Niveau(x) d'enseignement: 1 Numéro CDDB : 0478879
Centre d'intérêt: 540

Lapointe, Rita Eva.

L'observation systématique des élèves en classe / par Rita Eva Lapointe. -- [Montréal?] : R. E. Lapointe, 1988.
x, 167 p. ; 28 cm.

En tête du titre: Université de Montréal. "Mémoire présenté à la Faculté des études supérieures en vue de l'obtention du grade de maître ès arts (M. A.)". Bibliogr. : p. 164-167.

Descripteur(s): *Observation de la classe *Evaluation de l'étudiant *Mesure *Instrument de mesure *Primaire (niveau)

L'observation utilisée en classe par l'enseignant n'a pas acquis un statut reconnu de fiabilité. Cette recherche tente de rassembler, d'analyser et de présenter, sous forme de synthèse, un ensemble de questions et de réponses concernant l'observation systématique. Elle tente aussi d'adapter ces méthodes, conçues principalement en fonction des situations de recherche, à celles vécues en classe, dans le but de définir un moyen de mesure qui serait des plus utiles pour l'enseignant qui désire intervenir dans le développement des habiletés chez ses élèves.

Après avoir analysé la situation scolaire, l'auteur définit l'observation, ses objets et ses rôles. Elle présente ensuite un ensemble d'instruments qui peuvent être utilisés: anecdotes, rapports d'activités, grilles d'observation, échelles d'observation, protocole d'observation d'indices, protocole d'observation de catégories et divers types d'échelles d'appréciation. Elle analyse les facteurs qui déterminent la validité et la fidélité des observations. Elle décrit en détail trois instruments appropriés à l'observation en classe (grilles, échelles et protocole d'observation d'indices) et en donne des exemples d'utilisation. Elle propose enfin une stratégie d'observation applicable à la classe régulière.

Type de document: r

EDUQ 3470079

UNIVERSITÉ DE MONTRÉAL

L'OBSERVATION SYSTÉMATIQUE DES ÉLÈVES EN CLASSE

RITA-ÉVA LAPOINTE

AOÛT 1980

EDUQ 847007

UNIVERSITE DE MONTREAL

L'OBSERVATION SYSTEMATIQUE DES ELEVES EN CLASSE

PAR

RITA EVA LAPOINTE

SECTION MESURE ET EVALUATION

FACULTE DES SCIENCES DE L'EDUCATION

MEMOIRE PRESENTE A LA FACULTE DES ETUDES SUPERIEURES
EN VUE DE L'OBTENTION DU GRADE DE
MAITRE ES ARTS (M.A.)

AOÛT 1980

TABLE DES MATIERES

LISTE DES TABLEAUX vi

SOMMAIRE vii

INTRODUCTION ix

CHAPITRE

I. ANALYSE DE LA SITUATION 2

1. Les programmes d'étude 2

2. Les habiletés 4

3. L'organisation de la classe 7

4. Un premier ensemble de critères 11

5. L'objectif de la recherche 12

II. DEFINITIONS ET INSTRUMENTS 14

1. Une définition de l'observation 16

2. Les objets de l'observation 17

3. Les rôles de l'observateur 21

4. Les instruments d'observation 27

4.1 Les étapes de l'observation 28

4.2 Les résultats de l'observation 31

5. La présentation des instruments 36

5.1 Les anecdotes 37

5.2 Les rapports d'activités 40

5.3 Les grilles d'observation 44

5.4 Les échelles d'observation 46

5.5 Le protocole d'observation d'indices 48

5.6 Le protocole d'observation de catégories 51

5.7 Les échelles d'appréciation 53

5.7.1 Le rôle de l'observateur 54

5.7.2 Le type de résultats 55

5.7.3 L'échelle graphique descriptive . 59

5.7.4 L'échelle numérique ou alphabétique 60

5.7.5 L'échelle d'échantillons 62

5.7.6 L'utilisation des échelles d'appréciations 63

6. Une classification des instruments 64

III. VALIDITE ET FIDELITE 70

1. Les facteurs reliés à l'observateur 73

1.1 Le degré d'entraînement 73

1.2 La justesse de l'observateur 75

1.3 L'inférence 77

1.4 L'effet de halo 78

2. Les facteurs reliés à l'instrument 79

2.1 L'échantillonnage 79

2.2 Le cadre de référence 85

3. Les facteurs reliés à la séance d'observation 87

3.1 Les conditions de l'observation 87

3.2 Les effets observateur-observé 89

4. Les indices de fidélité 93

4.1 Le coefficient d'accord 94

4.2 Le coefficient de stabilité 96

4.3 Le coefficient de fidélité 99

5. La validité 101

IV. TROIS INSTRUMENTS ET QUELQUES EXEMPLES 106

1. Les grilles d'observation 107

2. Les échelles d'observation 110

3. Le protocole d'observation d'indices 116

4. Quelques exemples d'utilisation 125

4.1 Learning Accomplishment Profile 126

4.2 Cleveland Perceptual-Motor Rating Scale 130

4.3 Maryland Systematic Teacher Observation Instrument 133

4.4 Le protocole d'Anderson 137

V. PROPOSITION D'UNE STRATEGIE 144

1. La construction des instruments 147

2. L'encadrement de l'observateur 147

2.1 Le cadre de référence 147

2.2 L'entraînement de l'observateur 148

3. Les conditions de l'observation 151

3.1 L'échantillonnage un-zéro 152

3.2 L'échantillonnage instantané 155

TABLE DES MATIERES

CONCLUSION 160

REMERCIEMENTS 162

BIBLIOGRAPHIE 164

TABLEAU 1

Objets exigeant des procédures d'évaluation allant au-delà du test papier-crayon 20

TABLEAU 2

Une classification des instruments d'observation 68

TABLEAU 3

Les facteurs de fidélité et de validité 71

TABLEAU 4

L'ensemble des signes d'un protocole 120

TABLEAU 5

Protocole d'observation d'indices 121

TABLEAU 6

Extrait du Learning Accomplishment Profile 127

TABLEAU 7

Reproduction d'une fiche de consignation 139

SOMMAIRE

Ce mémoire étudie les problèmes reliés à l'observation en général et ceux liés à la situation particulière de l'observation d'élèves dans le cadre d'une classe régulière.

Après avoir défini l'observation et en avoir décrit les composantes, un ensemble d'instruments utilisés par les observateurs sont présentés. Les facteurs déterminant la validité et la fidélité des observations sont analysés et quelques critères et procédés pour assurer les qualités métrologiques des observations sont dégagés.

Avant cette étude théorique de l'observation, le cadre de l'observation d'élèves en classe est défini par l'analyse de la situation scolaire.

Trois instruments jugés les plus appropriés à cette situation sont présentés en détail et quelques exemples d'application sont considérés.

Enfin, le mémoire propose une stratégie d'observation applicable à la situation scolaire définie.

INTRODUCTION

INTRODUCTION

L'observation est souvent utilisée en classe par l'enseignant. Il s'en sert pour vérifier certains apprentissages, pour étudier certaines façons de faire, certaines habitudes de travail de ses élèves. Toutefois, ce n'est qu'exceptionnellement qu'on trouve une trace évidente de ces observations au bulletin scolaire des élèves.

L'observation en classe n'a pas acquis un statut égal aux instruments de mesure tels l'examen écrit ou l'examen oral. Que manque-t-il à l'observation de l'enseignant pour que ses résultats puissent être considérés fiables et valables au même titre que ceux des instruments de mesure généralement utilisés?

Lorsque l'observation est utilisée comme moyen de recueillir des informations, elle l'est dans des conditions bien précises; elle utilise des instruments qui sont élaborés selon des méthodes définies; elle cherche à obtenir des données fidèles et valides. Autant de sujets qui seront traités au cours de la présente étude.

Parallèlement à cette définition de l'observation systématique et de ses instruments, le désir de vérifier la possibilité que l'enseignant utilise ce type d'observation dans sa classe, se traduit d'abord par une analyse de la situation scolaire.

La comparaison constante des critères pratiques issus de cette situation avec ceux de la théorie de l'observation permet ensuite de percevoir des adaptations possibles et de proposer, à la fin du travail, une stratégie d'obervation appliquée à la classe régulière.

CHAPITRE PREMIER

ANALYSE DE LA SITUATION

La présente recherche se situe à la fois dans un cadre théorique et dans un cadre réel. C'est de ce dernier dont il sera question dans ce chapitre sur l'analyse de la situation. L'étude de l'observation concerne ici l'enseignant qui travaille auprès de groupes d'une trentaine d'élèves et qui est responsable de l'atteinte des objectifs prescrits par les programmes d'étude. L'analyse des diverses facettes de cette réalité scolaire permettra de dégager un ensemble de critères d'ordre pratique, propres à orienter les choix à faire.

1. Les programmes d'études

Certains programmes d'étude destinés aux élèves du primaire et du secondaire mettent l'accent sur le développement d'habiletés.

Le Programme d'étude du français, langue maternelle, au secondaire, par exemple, parlant de la langue maternelle comme d'un instrument de communication personnelle et sociale, indique que le "rôle de la classe de français ... est de développer chez l'élève l'habileté à utiliser efficacement cet instrument".[1]

---

1. Province de Québec, Ministère de l'Education, D.G.D.P., Programme d'étude - secondaire - Français langue maternelle, document no 16-3412-01, Québec, novembre 1979, p. 7

Ce programme définit des objectifs généraux concernant "les habiletés langagières nécessaires à la compréhension et à la production de discours oraux et écrits".[2] Par des objectifs terminaux pour chaque degré du secondaire, le programme propose ensuite un choix de discours que l'élève d'un degré donné doit pouvoir comprendre et produire, et réfère constamment aux éléments notionnels retenus pour ce degré. Cela constitue un portrait des habiletés langagières pour ce degré.

Ce Programme d'étude définit ensuite le rôle de l'évaluation comme étant de "déterminer dans quelle mesure les habiletés de l'élève coïncident avec ce portrait, s'en rapprochent ou s'en éloignent".[3]

Le Programme d'étude de la mathématique au primaire fournit un autre exemple de programme qui s'intéresse au développement d'habiletés.

---

2. Idem, p. 15

3. Idem, p. 86

Dans la présentation de ce programme, on retrouve l'objectif qui suit:

> "Ce programme vise donc, entre autres choses ... à promouvoir chez ce dernier (l'enfant) le développement de certaines habiletés intellectuelles, comme l'habileté à 'structurer' (établir des liens, organiser, classer, ordonner, ...), l'habileté à 'formaliser' (s'initier à une plus grande rigueur d'expression, à l'aide d'une terminologie et d'un symbolisme plus précis et mieux adapté), l'habileté à 'prouver' (à partir de faits admis, énoncer des conclusions sûres, par déduction ou par inclusion) ...".[4]

Le programme précise de plus que c'est parallèlement au développement de telles habiletés que l'enfant doit acquérir certains concepts mathématiques fondamentaux et maîtriser certaines techniques opératoires de base dans les domaines du nombre, de la géométrie et de la mesure.

2. Les habiletés

Ces programmes guident le maître dans son action auprès des élèves et ces derniers vivront de plus en plus de situations d'apprentissage visant le développement d'habiletés plutôt que l'acquisition de connaissances.

---

4. Province de Québec, Ministère de l'Education, D.G.D.P., Programme d'étude - primaire - Mathématique, document de travail no 16-2300-00, Québec, juin 1979, p. 10

Ces programmes établissent une différence entre le développement d'habiletés et l'acquisition de connaissances. La bibliographie du Programme d'étude de mathématique au primaire[5] cite des auteurs tels que Bloom, DeLandsheere, Harrow et Krathwohl. Cela laisse supposer que ces auteurs ont inspiré les rédacteurs du programme de mathématique lorsqu'ils ont défini les objectifs et aussi lorsqu'ils ont différencié habileté et connaissances.

Il est important, pour la présente discussion, de définir ces termes. Alors que connaissances est défini comme savoir et érudition sur divers sujets, habileté l'est comme qualité de celui qui est capable d'appliquer ce qu'il sait, de celui qui sait faire. Habileté est aussi défini comme le caractère de celui ou celle qui sait faire alors que connaissances fait référence à notion, idée de quelque chose.

---

5. Idem, p. 10

D'autres distinctions entre habiletés et connaissances apparaissent dans les taxonomies de Bloom,[6] Guilford,[7] Gagné et Merrill.[8]

Ainsi B.S. Bloom décrit les apprentissages en déterminant deux sous-ensembles: les objets de la connaissance et les habiletés intellectuelles ou processus mentaux utilisant les connaissances. Pour ces processus mentaux, il établit six catégories: la connaissance, la compréhension, l'application, l'analyse, la synthèse et l'évaluation.

On trouve une distinction correspondante entre habiletés et connaissances chez d'autres auteurs.

Ainsi, J.P. Guilford définit trois ensembles: les opérations, les contenus et les produits. Alors que ces derniers sont les résultats du traitement des informations, les opérations sont les processus de traitement et les contenus sont les informations ou connaissances à traiter au cours des opérations.

---

6. Gilbert DeLandsheere et V. DeLandsheere, Définir les objectifs de l'éducation, 2e éd., Paris, P.U.F., 1976, pp. 67-73

7. Idem, pp. 94-97

8. Idem, pp. 104-111

De même, R. Gagné et M.D. Merrill élaborant en collaboration une taxonomie de l'ensemble comportemental, définissent les comportements affectifs, psychomoteurs, et cognitifs illustrant des processus d'apprentissage. Les objets d'apprentissage sont tenus par ces deux auteurs, à l'extérieur de leur taxonomie. Cela permet de supposer qu'ils distinguent nettement les objets (que l'on pourrait ici aussi appeler connaissances), des comportements (que l'on pourrait appeler habiletés).

Il semble que ce soit les habiletés ainsi définies que l'école vise lorsqu'elle se donne pour objectif le développement d'habiletés. Les programmes qui guident de façon plus ou moins précise les enseignants comportent à la fois des objectifs concernant les connaissances et des objectifs concernant les habiletés telles que définies ci-haut.

3. L'organisation de la classe

Pour atteindre ces objectifs, l'école utilise ses ressources, compose avec des contraintes de divers ordres et s'engage dans une réflexion et une recherche pour améliorer son action. L'école tente d'améliorer l'environnement scolaire de l'élève et de lui proposer des situations propres à favoriser le développement de ses habiletés.

Le travail de l'enseignant consiste principalement à mettre tout en oeuvre en ce sens. Faire des exposés, fournir des exemples significatifs, diriger des travaux et des exercices, évaluer et critiquer les résultats sont quelques-unes des tâches de l'enseignant.

Choisissant d'étudier le développement des habiletés, on pourrait donc retrouver les situations suivantes dans les classes des écoles primaires ou des écoles secondaires.

Un groupe d'élèves apprend le mot juste pour désigner un objet, une action ou un sentiment. Il apprend aussi à utiliser ces mots au cours de différents exercices oraux. Il développe ainsi son habileté à utiliser le mot juste au cours d'une conversation.

Un groupe d'élèves apprend l'orthographe d'un ensemble de mots de toutes catégories. Dans diverses compositions ou rédactions, courtes ou longues, il saura l'orthographe d'usage des mots étudiés et de ceux rencontrés dans ses lectures.

Un groupe d'élèves a appris les résultats de la combinaison par addition, multiplication et division des nombres de 1 à 50. Il est engagé depuis quelque temps dans une série d'exercices de calcul oral appliqué à la vie quotidienne.

Un groupe aborde la division avec un nombre de trois chiffres au diviseur. Avant de passer à l'entraînement, il lui faut s'assurer qu'il possède la bonne façon de procéder.

Un groupe d'élèves a préparé une série d'arguments qui appuient la position qu'il doit défendre. Lors de la discussion, il exploite ces arguments en tenant compte du climat de la discussion et du développement de l'argumentation adverse.

Un groupe analyse une situation problématique dans leur école. Ils ont rassemblé des informations sur la situation actuelle et sur la situation souhaitée par les différents groupes de personnes impliqués: élèves, représentants d'élèves, professeurs du département. Le groupe cherche à établir une liste de solutions possibles.

Dans chacune des situations qui précèdent, l'enseignant voudra compléter son action en évaluant et critiquant les résultats obtenus par les élèves. Il voudra mesurer et évaluer les apprentissages à deux moments principaux: pendant l'apprentissage, pour ajuster son enseignement aux résultats obtenus, et à la fin de périodes d'apprentissage fixées par le calendrier scolaire, pour faire le point et faire rapport aux diverses personnes intéressées, entre autres, l'élève et ses parents.

A la fin des temps fixés par le calendrier scolaire, l'enseignant pourra classer ou noter les résultats d'un élève selon le degré comparatif d'habileté qu'il a démontré au cours d'un occasion ou d'une épreuve donnée. L'enseignant utilisera à cet effet un ensemble de critères plus ou moins objectifs, plus ou moins personnels.

4. Les besoins de l'enseignant

L'enseignant qui veut mesurer les habiletés en cours d'apprentissage dans le but d'orienter ce même apprentissage a besoin de savoir de façon précise où sont les failles dans le processus suivi par les élèves lors du recours à une habileté donnée. Des informations sur le seul résultat terminal de l'utilisation de cette habileté ne sauraient lui suffire puisqu'il veut agir à divers moments du développement de l'habileté.
Le maître doit plutôt recueillir des informations sur les diverses manifestations échelonnées dans le temps et rassemblées dans un processus correspondant à l'habileté en développement, habileté plus ou moins intégrée par l'élève.

Ce dont l'enseignant a besoin, ce sont des procédures et des instruments:

- permettant de recueillir des informations précises
- visant le déroulement d'un processus (et non les résultats de celui-ci)
- fournissant des informations pour chacun de sa trentaine d'élèves
- pouvant s'utiliser pour 2 à 5 élèves à la fois
- n'exigeant pas une somme énorme de travail de compilation

. ne consommant pas trop de temps en classe lors de leur application

. étant d'utilisation simple

. produisant des données valables et fiables

Or, l'observation est une façon de recueillir des informations sur le déroulement d'un processus aussi bien que sur les résultats obtenus à la fin du processus.

5. L'objectif de la recherche

Pour arriver à ses fins, celui qui enseigne à un groupe d'élèves pourrait-il utiliser l'observation? A quelles conditions l'observation pourrait-elle fournir, dans le contexte d'une classe, une stratégie pratique pour recueillir des informations valables et fiables?

Ce sont là les questions auxquelles les pages qui suivent cherchent à répondre.

CHAPITRE II

DEFINITIONS ET INSTRUMENTS

Dans le but de définir le cadre théorique dans lequel s'insère l'observation, une revue de littérature a été effectuée sur le sujet. Une synthèse des informations recueillies est présentée dans les pages qui suivent.

Les écrits répertoriés sont principalement des articles de revues scientifiques, articles publiés depuis 1953. Quelques chapitres de livres spécialisés en mesure et évaluation en éducation, en psychologie et en sociologie, se sont également révélés pertinents.

L'analyse des informations rassemblées lors de cette revue de littérature a d'abord consisté à dégager les éléments particuliers à chaque auteur puis à considérer les liens et relations possibles entre ces divers éléments. La recherche des principes d'organisation de ces éléments et des liens qui se tracent entre eux a permis ensuite l'élaboration d'une classification des instruments d'observation et d'un ensemble de critères assurant la valeur des observations faites.

La synthèse des informations comporte trois volets. Le premier propose d'abord une définition de l'observation, une définition des objets et des rôles possibles de l'observateur. Puis, l'ensemble des procédures et instruments d'observation relevés est présenté et une classification est proposée. Lors de la présentation des instruments, les caractéristiques de chaque groupe sont confrontées aux besoins définis par la situation problématique et les instruments retenus pour étude subséquente sont identifiés.

Le second volet relève les questions soulevées par les qualités métrologiques de validité et de fidélité des résultats de l'observation, en rapport avec les instruments étudiés. Ce sera l'objet du troisième chapitre.

Dans un dernier temps, les instruments retenus sont analysés de façon précise en regard de la démarche ou procédure qu'ils proposent ainsi que des façons dont ils garantissent la qualité des résultats obtenus par leur utilisation. Ce volet, qui constitue le quatrième chapitre, comporte en plus une étude de quelques exemples d'utilisation de l'observation, exemples qui permettent l'étude de la praticabilité de cette utilisation, même si leur cadre n'est pas celui d'une classe.

1. Une définition de l'observation

Pour identifier les caractéristiques de l'observation, on pourrait la comparer à l'examen. Mais en procédant ainsi, on supposerait déjà qu'utilisée en milieu scolaire, l'observation puisse avoir les mêmes fonctions que l'examen: mesurer les apprentissages.

Avant d'étudier cette possibilité, la définition habituelle de l'observation contribuera à en dégager les diverses composantes.

> . Action de considérer avec attention les choses physiques ou les choses morales.
>
> . Procédé logique à l'aide duquel on constate toutes les particularités du phénomène en lui-même, sans le troubler par l'expérimentation.

L'observation est donc un procédé, une manière de faire pour constater, puis consigner ou noter si on le désire, ce qui rend un phénomène différent d'un autre. L'observation cherche à ne pas troubler le phénomène observé, à ne pas l'orienter par des questions, des commentaires ou des directives au cours de son déroulement.

Sorte d'interrogatoire que l'on fait subir à quelqu'un, l'examen tente aussi de faire ressortir les différences. La série de questions qui le constitue oriente le travail de l'examiné selon un plan déterminé par celui qui a construit l'examen.

On peut donc dire que les deux procédés, l'observation et l'examen, visent à établir les différences qu'il y a entre deux choses. Ce qui caractérise l'un ou l'autre, c'est le procédé utilisé. L'observation constate en intervenant le moins possible alors que l'examen questionne et donne des directives.

2. Les objets de l'observation

Si les procédés de l'observation sont différents de ceux de l'examen, ses objets le sont-ils? Peut-on choisir indifféremment l'observation ou l'examen pour recueillir des informations sur les différences et les particularités de phénomènes?

Par exemple, lorsqu'on désire des informations sur l'utilisation d'une technique, utilisera-t-on l'examen ou l'observation?

Il faudrait d'abord savoir si l'utilisation de cette technique donne nécessairement un type particulier de résultats. Si tel est le cas, on pourra construire un examen autour de l'objet de mesure qu'est ici la technique utilisée, et, par l'étude du résultat obtenu par l'élève, conclure au sujet de l'habileté de l'élève à utiliser cette technique.

Par contre, si on arrive à la conclusion contraire que la technique que l'élève doit maîtriser ne donne pas nécessairement un type particulier de résultats, il faudra observer l'élève pendant qu'il travaille pour savoir s'il utilise ou non cette technique, et s'il l'utilise adéquatement. De plus, selon le degré de précision de l'instrument d'observation utilisé, les données recueillies au sujet de la méthode de travail de l'élève permettront d'intervenir de façon plus ou moins précise auprès de cet élève, après la session d'observation.

Dans son livre Measurement and Evaluation in Teaching, Norman E. Gronlund[1] consacre deux chapitres aux techniques d'observation. Des auteurs étudiés, il est le seul à présenter un ensemble d'objets demandant qu'on emploie l'observation.

Le tableau qui suit est une traduction du tableau qu'il présente.

---

1. Norman E. Gronlund, Measurement and Evaluation in Teaching, 3e éd., New-York, Macmillan, 1976, 590 p.

Tableau 1

Objets exigeant des procédures d'évaluation allant au-delà du test papier-crayon[a]

Objet	Comportements
Habiletés	Parler, écrire ... Habiletés de travail, habiletés d'étude, habiletés sociales ...
Habitudes de travail	Planification adéquate, utilisation du temps, de l'équipement et des ressources; la démonstration de caractéristiques telles: initiative, créativité, persistance, fiabilité ...
Attitudes sociales	Attention au bien être des autres, respect des lois ...
Attitude scientifiques	Ouverture d'esprit, objectivité, sensibilité aux relations de cause à effet, esprit de recherche, ...
Intérêts	Sentiments exprimés en regard de diverses activités d'ordre éducatif, mécanique, esthétique, scientifique, social, récréatif, professionnel ...
Appréciations	Sentiments de satisfaction et de joie envers la nature, la musique ...
Adaptations	Relations avec les pairs; réactions aux critiques, aux éloges, à l'autorité; stabilité émotionnelle, adaptabilité sociale.

a. Norman E. Gronlund, Measurement and Evaluation in Teaching, 3e éd., New-York, Macmillan, 1976, p. 308

3. Les rôles de l'observateur

Une autre caractéristique de l'observation c'est la diversité des rôles que peut tenir l'observateur.

Comme le mentionne Heyns et Zander "The use of observers means the use of people as measuring instruments."[2] Contrairement à l'examen ce n'est pas un ensemble de questions qui fournira les données recherchées, c'est l'observateur lui-même qui, orienté ou non par ce qu'on apppellera l'instrument d'observation, rassemblera les données recherchées.

L'observateur peut se situer à des points de vue différents et peut accomplir différentes tâches. Selon le point de vue qu'il adoptera et la tâche qu'il aura à accomplir, le jugement et l'expérience de l'observateur entrera plus ou moins en cause.

---

2. Heyns, R. et R. Lippitt, Systematic Observational Techniques, in G. Lindzey (Ed.), "Handbook of social psychology", Vol. 1, Cambridge, Mass., Addison-Wesley, 1954, p. 416

Ainsi, Dowley [3] propose que l'observateur peut prendre note:

a) des actions des élèves,

b) des sentiments des élèves à propos de ce qu'ils font, ou

c) de ses propres impressions et interprétations de a et b.

Cette classification des observations en trois niveaux ne rencontrent pas la classification que fait l'ensemble des auteurs lus des tâches de l'observateur. En effet, ceux-ci regroupent a et b en un seul type, l'observation proprement dite et parlent de c en terme d'évaluation (rating). Cependant, la classification de Dowley laisse entrevoir la différence de point de vue entre l'observateur qui consigne des faits et des actions et celui qui interprète, évalue et cote ces mêmes faits et actions. Egalement, en divisant les actions des élèves (a) et les sentiments de ceux-ci à propos de leurs actions (b), Dowley veut indiquer à l'observateur qu'il lui faudra être vigilant afin d'isoler ou non ses observations du contexte qu'il percevra comme tel, selon le plan qu'il aura établi.

---

3. Edith M. Dowley, Cues of Observing Children's Behavior, in "Childhood Education", vol. 45, no 9, mai 1969, pp. 517-521

On peut dire que Dowley rappelle à l'observateur qu'il est une personne devenue instrument de mesure, une personne qui doit tendre à être consciente de ses perceptions, de ses impressions et de ses opinions lorsqu'elle agit en tant qu'observateur. En ce sens, Dowley rejoint Heyns et Zander cités plus haut.

Dans les faits, le déroulement d'une observation justifie les mises en garde faites par ces auteurs. L'analyse de chacune des étapes d'une observation montre l'observateur utilisant ses perceptions et son jugement. Ainsi, l'observateur doit:

1. considérer un objet, une action, un comportement;
2. définir ou décrire ce qu'il perçoit;
3. comparer sa définition à celle fournie dans son plan d'observation;
4. juger son observation et enfin;
5. consigner cette observation selon les règles établies par son plan d'observation.

Par ailleurs, ce ne sont pas tous les types d'observation qui exigent que l'observateur effectue les quatre dernières étapes décrites, et l'étape de catégorisation peut faire plus ou moins appel au jugement et à l'expérience des observateurs.

Ainsi, l'observateur qui rédige un rapport d'activité (self-report) effectuera les étapes 1, 2 et 5 seulement alors que s'il est engagé dans une observation systématique avec un plan d'observation définissant, par exemple, huit catégories, l'observateur devra effectuer les cinq étapes décrites à chacune de ses observations. C'est à cette dernière situation que se rattache le commentaire suivant de Kerlinger sur la tâche principale de l'observateur: "The fundamental pratical job of the observer is to assign behaviors to categories".[4]

Les cinq étapes d'une observation son également nécessaires lorsqu'on demande à l'observateur de catégoriser son observation non pas en fonction de catégories reliées à l'objet observé mais plutôt en fonction de catégories qualitatives, du plus au moins, de l'excellent au médiocre. Ici, l'observateur pourrait changer son titre pour celui d'évaluateur car son plan d'observation repose sur l'utilisation d'une échelle d'appréciation (rating scale).

---

4. F.N. Kerlinger, Foundations of Behavioral Research, New-York, H.R.W., 2e éd., 1973, p. 540

Quelques auteurs indiquent cependant que les résultats obtenus par le biais des échelles d'appréciation sont des résultats quantitatifs au même titre que ceux obtenus par le biais des instruments d'observation (check list, scaled check list, category system, sign system). Ainsi, Ghiselli et Brown écrivent: "Ratings are procedures for quantifying the opinions of one person about the behavior of another".[5]

En fait, il semble que le terme "rating" devrait plutôt être traduit par le terme appréciation. Lorsque le terme "rating" est rendu par évaluation cela apporte de la confusion étant donné que ce mot recouvre une réalité plus large et qu'il est difficile de penser que cette évaluation est quantification. "Rating scale" deviendrait donc grille d'appréciation ou échelle d'appréciation.

---

5. E.E. Ghiselli et C.W. Brown, Personnel and Insdustrial Psychology, New-Yok, McGraw-Hill, 1955, p. 88

L'utilisation de catégories qualitatives, demande à l'observateur de faire appel à son expérience et à son jugement. Ainsi que l'explique Travers (1969):

> "A rater is employed to derive whatever signifiance may be possible from his observations, and this he does on the basis of judgment, which in turn is based o' experience".[6]

Dans cette situation, l'observateur tient compte d'un ensemble d'informations dont il fait une synthèse avant de choisir la catégorie qualitative qui décrira son opinion. C'est ce qu'indique Travers lorsqu'il écrit: "Rating is a way by which numerous events that are observed are somehow summarized and combined".[7]

La dernière situation décrite par Travers montre un observateur très impliqué en tant que personne dans la procédure qui résultera en un ensemble d'appréciations. Les situations décrites plus avant demandaient moins d'implication de la part de l'observateur. Mais quelle que soit la situation, l'observateur est un élément essentiel de l'observation.

---

6. Robert M.W. Travers, An Introduction to Educational Research, New-York, Macmillan, 1958, p. 125

7. Idem, p. 143

4. Les instruments d'observation

L'observation, définie dans la section précédente comme procédé ou manière de faire pour constater, puis consigner ou noter si on le désire, ce qui rend un phénomène différent d'un autre, prend diverses formes. Cette diversité apparaît soit dans le procédé même, soit dans l'instrument utilisé pour consigner les résultats de l'observation.

Les auteurs répertoriés présentent plusieurs types d'instruments et de procédés. Chacun d'eux sera étudié ci-après dans leur forme, le rôle qu'il donne à l'observateur et le type de résultats obtenus.

Avant cette étude, la présentation d'une vue d'ensemble des étapes possibles de la démarche de l'observateur lorsqu'il utilise un instrument d'observation de même qu'un commentaire autour de la question des résultats de l'observation, permettront de mieux faire ressortir les particularités de chaque instrument qui sera décrit dans les pages qui suivront.

4.1 Les étapes de l'observation

Une synthèse générale des informations recueillies au cours de la revue de littérature effectuée, permet de dire que l'observateur est habituellement orienté par un instrument, plus ou moins structuré, plus ou moins précis. Sa façon de procéder sera conséquemment plus ou moins structurée et plus ou moins précise.

Cette même synthèse nous amène à constater qu'au plus, la démarche de l'observateur peut comporter les cinq étapes déjà mentionnées: attention, description, comparaison, jugement et consignation.

Pour plus de précision, chaque étape peut être définie comme suit:

Première étape: diriger son attention -- ici, l'observateur se met en situation d'écoute, d'attention globale, afin de reconnaître des signes, des indices. L'observateur considère une situation, un objet, un procédé, et décide de diriger son attention sur l'un ou l'autre de ces éléments pour être à même de percevoir le plus de choses possibles.

Deuxième étape: définir ou décrire ses perceptions -- l'observateur, pendant cette étape, transforme ses perceptions en quelque chose de communicable.

a) S'il s'agit d'un instrument d'observation seul, il décrit ce qu'il perçoit de façon consciente, principalement par la vue, l'ouie et le toucher.

b) S'il s'agit d'un instrument d'appréciation, il précise son opinion, son appréciation des observations qu'il a faites.

Troisième étape: comparer sa définition à celle fournie -- lorsque l'observateur utilise un instrument pour consigner ses observations, c'est au cours de cette étape de comparaison entre sa description et celle donnée par l'instrument, qu'il trouvera l'information qui lui est nécessaire, concernant les écarts entre sa définition et l'une ou l'autre des définitions proposées, pour procéder à l'étape suivante.

Quatrième étape: juger son observation -- cette étape prend une allure différente selon le type d'instrument d'observation. L'observateur, connaissant par l'étape précédente, les écarts plus ou moins grands entre sa définition et l'une ou l'autre de celle proposée par son instrument d'observation, prendra une décision:

a) dans certains cas, il décidera si oui ou non l'élément défini est présent;

b) dans d'autres cas, il décidera de classer son observation, dans la catégorie de faits qui réduira au maximum les écarts de définition;

c) dans d'autres cas, il décidera de classer son appréciation en regard de la catégorie d'appréciations qui minimisera les écarts de description.

Cinquième étape: consigner son observation selon les règles établies dans son plan d'observation -- à cette étape, l'observateur doit maintenant faire une certaine marque à l'endroit prévu sur un formulaire déterminé. Ce pourra être une croix, un tiret, un chevron ou autre. L'observateur doit suivre en cela les règles établies.

4.2 Les résultats de l'observation

Selon la problématique définie au début du présent rapport, la procédure ou l'instrument d'observation qui sera retenu pour répondre aux besoins mentionnés, devra fournir à l'enseignant des données, des résultats fiables et valables. Ces qualificatifs font appel aux qualités de toute mesure.

Les résultats de l'observation seraient-ils des résultats de mesure? Serait-ce mesurer que d'observer?

Quelques-uns des auteurs répertoriés se sont posés cette même question. Stevens étudiant les relations entre la mathématique, la mesure et la psychophysique, répondait par l'affirmative:

> "In its broadest sense, measurement is the assignment of numerals to objects or events, according to rules".[8]

A la suite d'une observation, on peut au moins dire si oui ou non l'élément observé était présent, et on peut assigner une valeur numérique de 1 au oui et de 0 au non selon une règle pré-établie. On aurait alors un résultat de mesure, au sens large, tel que défini par Stevens.

---

8. S. Stevens, Mathematics, Measurement and Psychophysics, in S. Stevens (Ed.), "Handbook of Experimental Psychology", New-York, Wiley, 1951, p. 1

C'est là le point de départ de Kerlinger qui lui-même cite Stevens et ajoute, dans son livre sur les fondements de la recherche sur le comportement:

> "The social scientist must also observe human behavior ... He treats the observation of behavior as a part of a measurement procedure: he assigns numerals to objects, in this case human behavioral acts or sequences of acts, according to rules".[9]

Delgado et Delgado[10] dans un article traitant d'une approche objective de la mesure du comportement, identifie quatre aspects au problème de la mesure:

a) isoler des variables;

b) établir et définir des unités de mesure;

c) comparer les unités;

d) interpréter les résultats des comparaisons.

La démarche de l'observateur décrite plus haut assure que celui-ci effectue des actions qui se rattachent directement aux aspects a, c et d identifiés. Le deuxième aspect, est souvent pris en charge par la procédure ou l'instrument d'observation qui sera choisi et utilisé par l'observateur.

---

9. F.N. Kerlinger, op. cit., p. 537

10. R.R. Delgado et J.M.R. Delgado, An Objective Approach to Measurement of Behavior, in "Philosophy of Science", vol. 29, juillet 1962, p. 540

Puisqu'elle apporte une réponse à chacun des quatre aspects du problème de la mesure, l'observation a le droit de se réclamer de celle-ci et l'on peut dire que les résultats obtenus par l'observation sont des résultats de mesure. Toutefois, la qualité des résultats dépendront de la qualité des réponses apportées aux problèmes soulevés par l'observation, entre autres, l'entraînement de l'observateur, l'inférence, l'effet de halo, l'échantillonnage et les conditions de l'observation.

Enfin, la définition que donne Cronbach (1970) d'un test, renforce l'affirmation qu'observer, selon certaines règles, c'est mesurer. Pour lui, un test c'est:

> "a systematic procedure for observing a persons's behavior and describing it with the aid of a numerical scale or a category system".[11]

On peut lire cette définition en pensant à un test standardisé et la trouver pertinente. On peut la relire en pensant à certains instruments d'observation et la trouver également pertinente.

---

11. L.J. Cronbach, Essentials of Psychological Testing, New-York, Harper & Row, 1970, p. 26

Les résultats de mesure peuvent appartenir à l'une des quatre échelles de mesure suivantes:

- l'échelle nominale,
- l'échelle ordinale,
- l'échelle d'intervalles,
- l'échelle de rapports.

L'appartenance à l'une ou à l'autre de ces échelles est fonction du nombre de propriétés communes que les résultats obtenus ont avec les nombres, ces propriétés étant l'identité, l'ordre et l'additivité. Lorsque les résultats de mesure possèdent la propriété d'additivité, propriété fondamentale des nombres, il est possible d'effectuer les quatre opérations (addition, soustraction, multiplication et division) sur ces résultats au cours de leur compilation et de leur interprétation.

La considération de la façon dont les résultats sont obtenus par chaque instrument d'observation étudié ci-après, permettra de déterminer l'échelle de mesure leur correspondant et par conséquent le type d'interprétation qu'il sera possible de faire.

Dans le processus d'observation, à la deuxième étape, l'observateur définit ses perceptions. De telles définitions peuvent par la suite être classées, d'après des définitions ou des catégories données par l'instrument d'observation utilisé.

La classification se trouve à la base de toute mesure. De la qualité de cette classification dépend la qualité de la mesure et de ses résultats. Guilford[12] affirme qu'il est important que les classes ou catégories soient 1) bien définies, 2) mutuellement exclusives (si possible), et 3) exhaustives.

Il faut noter que le terme classe est pris ici dans un sens très large. Il n'est pas nécessaire que plusieurs éléments puissent être rassemblés sous un seul titre pour que ce titre porte le nom de classe. On peut ne trouver qu'un seul élément par classe si cet élément est une classe par lui-même c'est-à-dire qu'il a une identité propre, qu'il exclut les autres éléments des semblables et comprend tous les éléments semblables à lui-même. Cet élément constituera alors une classe ou catégorie au sens métrique du terme.

---

12. J.P. Guilford and B. Fruchter, Fundamental Statistics in Psychology and Education, New-York, McGraw-Hill, 1978, p. 15

5. La présentation des instruments

Les instruments étudiés dans la présente recherche seront présentés dans l'ordre suivant:

5.1 Les anecdotes (anecdotal records)

5.2 Les rapports d'activité (self-report technique)

5.3 Les grilles d'observation (checklists)

5.4 Les échelles d'observation (scaled checklists)

5.5 Le protocole d'observation d'indices (sign system)

5.6 Le protocole d'observation de catégorie (category system)

5.7 Les échelles d'appréciation

a) l'échelle graphique descriptive

b) l'échelle d'appréciation numérique ou alphabétique

c) l'échelle d'échantillons

5.1 Les anecdotes (anecdotal records)

Ce procédé d'observation consiste en une description factuelle d'incidents ou d'événements jugés significatifs par l'observateur. L'acception usuelle du terme anecdote ne demande pas qu'un fait soit significatif pour devenir anecdotique; il peut être d'aspect secondaire ou sans portée, en autant qu'il soit intéressant. Retenir le terme anecdote comme l'on fait tous les traducteurs c'est ajouter à sa définition habituelle.

L'observateur, ayant une idée générale de ce qu'il recherche, décide de prendre note d'un ou de plusieurs faits qu'il juge appropriés, sous forme d'un écrit circonstancié. La rédaction de l'anecdote a lieu le moins de temps possible après le fait.

L'observateur qui utilise ce procédé de façon régulière, produira avec le temps, des dossiers anecdotiques.

Aucun des auteurs répertoriés qui étudient ce procédé[13,14], ne mentionne l'utilisation d'un formulaire ou d'un instrument quelconque, ni d'un plan d'observation précis à être suivi par l'observateur.

---

13. N.E. Gronlund, op., cit.

14. W.J. Popham, Educational Evaluation, Englewood Cliffs, Prentice Hall, 1975, 328 p.

Dans l'utilisation de ce procédé, l'observateur accomplit les 1re, 2e et 5e étapes d'une observation.

L'observateur (1) est en état d'éveil de façon globale. Lorsqu'il juge qu'un incident est significatif, il y porte une attention particulière et (2) le garde en mémoire avec le plus de détails possible. Dès qu'il en a l'occasion, (5) il rédige un compte rendu de l'événement (l'anecdote) cherchant à noter séparément les faits et l'interprétation qu'il en donne.

De ce type d'observation résulte un ensemble de données, comprenant les faits et leur interprétation, qui peuvent se révéler fort utiles pour la compréhension d'une situation.

Il est nécessaire, note Gronlund[15], de limiter les observations et les comptes rendus par un plan prédéterminé visant des types précis de comportements et visant quelques élèves à la fois. Vouloir produire des comptes rendus anecdotiques pour tous les élèves d'un groupe d'une trentaine, et à propos d'un grand nombre de comportements, faits ou événements, rendrait la procédure impraticable.

---

15. N.E. Gronlund, op. cit., pp. 309-314

Au rang des avantages de cette procédure d'observation, Popham la comparant avec les procédures de l'observation systématique conclut:

> "The observer is free to note any behavior that appears important and is not required to focus on the same traits for all subjects. ... (This procedure) can provide a rich change of pace from the bland quantitative orientation of most observation procedures".[16]

Cette richesse dont parle Popham résiderait surtout dans le fait, comme le souligne N.E. Gronlund[17], que le compte rendu anecdotique fournit des descriptions qui tiennent compte des situations naturelles entourant ou provoquant l'événement. Cette procédure permet aussi de tenir compte d'événements exceptionnels mais non moins significatifs et tend à augmenter l'attention et la compréhension (awareness) de l'observateur.

Utiliser le compte rendu anecdotique comme système d'observation est presque impensable vu le temps qu'il faudrait consacrer pour maintenir un tel système. A cet inconvénient, N.E. Gronlund[18] ajoute qu'il est difficile d'être objectif lorsqu'on décrit un événement et que cette procédure ne produit pas un échantillon adéquat du comportement de la personne observée.

---

16. W.J. Popham, op. cit., p. 98
17. N.E. Gronlund, op. cit., p. 311
18. Idem

Bien que N.E. Gronlund[19] propose des moyens d'augmenter l'efficacité des comptes rendus anecdotiques, il apparaît que pour recueillir systématiquement des informations au sujet du développement des habiletés pour un groupe d'une trentaine d'élèves, il ne faut pas songer à recourir à cette procédure.

5.2 Les rapports d'activité (self-report technique)

Gronlund[20] mentionne des moyens de recueillir des informations à propos de comportements passés ou bien de la vie intérieure d'une personne, opinions, goûts et intérêts, informations qui peuvent s'avérer utiles à l'éducateur. Ces moyens sont l'entrevue, les listes de pointage d'activités ou de problèmes (activity checklists, problem checklists), les inventaires de personnalité, d'intérêts, et les échelles d'attitudes.

De ces divers moyens l'entrevue transformée en rapport d'activité offre une formule intéressante pour la situation scolaire où l'on désire des informations sur les apprentissages d'une méthodologie.

---

19. Idem, p

20. Idem

Le rapport d'activité consisterait en la rédaction par l'élève, d'un compte rendu, plus ou moins élaboré, décrivant l'activité qu'il a menée, la démarche suivie, et, s'il y a lieu, une critique des objectifs, de la démarche utilisée et des résultats de cette activité. On pourrait fournir à l'élève un plan pour la rédaction de son rapport.

L'entrevue individuelle pourrait apporter les mêmes informations mais le temps que devrait y consacrer l'enseignant serait beaucoup plus grand.

Dans cette procédure d'observation, il peut y avoir deux observateurs: d'abord l'élève. En effet, celui-ci doit prendre distance par rapport aux objectifs qu'il s'était fixé, à la démarche qu'il a utilisé et aux résultats obtenus. Il doit ensuite rédiger le rapport d'activité selon les règles qu'il établira lui-même ou celles que l'enseignant aura données.

Le deuxième observateur, l'enseignant, pourra alors noter ou apprécier ce rapport en utilisant soit une grille d'observation, soit une échelle d'observation ou bien, l'une ou l'autre des échelles d'appréciation décrites plus loin.

Pour utiliser cette procédure de cette façon, avec deux observateurs, il est nécessaire que l'élève soit suffisamment âgé et soit capable de rédiger lui-même le rapport d'activité.

Un deuxième mode d'utilisation du rapport d'activité, serait de le faire rédiger par une tierce personne, ni l'élève, ni l'enseignant. Ce pourrait être un élève plus âgé, ou un observateur externe. Dans ce cas, il s'agirait d'une activité de courte durée, alors que pour le premier mode d'utilisation, l'activité peut s'étendre sur une longue période de temps.

Contrairement à la procédure d'entrevue, dont le rapport d'activité s'inspire, la procédure de celui-ci se réalise en deux temps distincts. Le premier, rédaction du rapport, est une procédure qui lui est propre. Le deuxième temps, notation ou appréciation du rapport, fait appel à des procédures qui seront discutés dans les pages à venir.

L'observateur qui rédige le rapport d'activité, que ce soit la personne même qui a fait l'activité ou quelqu'un d'autre, effectue, comme pour le compte rendu d'une anecdote, les 1re, 2e et 5e étapes d'une observation.

En effet, cet observateur (1) considère l'activité dans son ensemble et choisit de faire rapport au sujet d'événements qu'il juge significatifs.(2) Il prend alors note de détails qu'il juge pertinents en fonction de l'activité et en fonction du plan du rapport qu'il veut produire. (5) Lorsque l'activité est terminée, il rédige le rapport selon certaines règles données, ou, à défaut, selon celles qu'il a établies lui-même.

Le rapport d'activité peut fournir à l'enseignant un large ensemble de renseignements lui permettant de connaître, comprendre, analyser et critiquer non seulement les objectifs et les résultats d'une activité mais aussi la démarche et la méthode de travail de ses élèves. La validité et la fiabilité des renseignements rassemblés, reposent pour une grande part sur l'observateur qui rédige le rapport d'activité. La part de validité qui s'appuie sur la pertinence de ces renseignements peut, par contre, être assez bien garantie si l'enseignant fournit à l'élève un plan d'observation faisant aussi office de plan du rapport.

Par ailleurs, il apparaît que cet instrument d'observation soit mieux adapté aux activités faisant la synthèse de plusieurs habiletés, qu'à celles faisant appel à une ou deux habiletés en développement. Comme il s'agit, dans la présente recherche, de cette dernière situation, le rapport d'activité ne fera pas l'objet d'une étude plus approfondie.

5.3 Les grilles d'observation (checklists)

Les grilles d'observation, appelées listes de vérification par Morissette[21] et liste de pointage par D. Bélanger[22] dans sa traduction de Selltiz, Wrightsman et Cook présentent à l'observateur une liste de caractéristiques ou d'actions à côté desquelles un espace est prévu pour que l'observateur indique, par une marque simple, la présence ou l'absence de la caractéristique recherchée, la réalisation ou l'omission d'une action de la suite de celles attendues.

---

21. Dominique Morissette, Les examens de rendement scolaire, Québec, P.U.L., 1979

22. Claire Selltiz, L.S. Wrightsman, et al ..., Les méthodes de recherche en sciences sociales, Montréal, H.R.W., 1977 (Traduit par David Bélanger)

La grille d'observation est en quelque sorte un plan d'observation pour un événement ou autre objet d'observation. Son utilisation suppose que l'on a prévu faire une observation et que l'on a choisi ou préparé un instrument à cette fin.

Les listes d'actions ou de caractéristiques qui constituent la partie la plus importante des grilles d'observation, peuvent varier en longueur et en complexité.

L'utilisateur des grilles d'observation réalise toutes les étapes d'une observation. (1) Il dirige son attention sur l'objet à observer, (2) décrit ce qu'il perçoit, (3) compare sa description à celle fournie par la grille d'observation qu'il utilise, (4) juge son observation et décide s'il la classera dans le oui ou le non, et (5) consigne celle-ci suivant les règles établies.

Les résultats obtenus par la grille d'observation appartiennent à l'échelle de mesure ordinale. Les observations sont divisées en deux classes: présent ou absent, oui ou non. Pour obtenir le résultat de chaque objet observés on compte le nombre de oui. Il est possible ensuite d'ordonner les objets observés selon leurs résultats.

Lorsque confrontée aux critères définis par la situation problématique à la base de cette recherche, la grille d'observation semble mériter une étude plus poussée. Tout indique à ce stade-ci que la grille d'observation pourrait être appropriée pour les besoins décrits.

5.4 Les échelles d'observation (scaled checklists)

Apparentée aux grilles d'observation, les échelles d'observations sont présentées par Ghiselli et Brown[23]. Elles ont l'avantage sur les premières de fournir à l'observateur un résultat final pondéré.

Tout comme la grille, l'échelle d'observation est constituée d'un assemblage d'énoncés de caractéristiques recherchées ou d'actions attendues. La différence entre les deux tient dans le fait qu'au cours d'une étape de l'élaboration de l'échelle, on a attribué une valeur chiffrée à chaque énoncé.

---

23. E.E. Ghiselli et C.W. Brown, op. cit., p. 110

La procédure d'utilisation de l'échelle d'observation demande à l'observateur d'indiquer par une marque simple, la présence des caractéristiques recherchées ou la réalisation d'une action de la série d'actions attendues. L'observateur peut marquer ainsi autant d'énoncés qu'il juge à propos.

Au terme de la session d'observation, l'objet ou l'événement observé recueille la moyenne des valeurs attribuées aux items marqués par l'observateur.

Comme pour la grille d'observation, l'observateur complète les cinq étapes d'une observation. La décision qu'il prendra à la quatrième étape sera celle de présence ou d'absence de l'élément recherché.

Les résultats obtenus par l'utilisation de l'échelle d'observation appartiennent à l'échelle de mesure à intervalles. Une valeur a été attribuée à chaque caractéristique par le biais de la technique des intervalles apparemment égaux de Thurstone et le résultat est constitué de la moyenne des valeurs attribuées aux caractéristiques jugées présentes par l'observateur.

A première vue, cet instrument d'observation semble offrir une solution intéressante à la problématique posée au départ. Cette procédure sera donc analysée plus en détails dans la section suivante.

5.5 Le protocole d'observation d'indices (sign system)

Pour la recherche sur l'enseignement, le climat d'une classe, les relations maître-élève et autres sujets de recherche, deux méthodes d'observation très élaborées ont été développées. Ces méthodes indiquent de façon très explicite ce qu'il faut observer, ce qu'il faut noter, quand et comment il faut le faire. Nous croyons que cela justifie le terme protocole d'observation. Dans un cas, il s'agira d'observation d'indices, dans l'autre, d'observation de catégories.

Le protocole d'observation d'indices, comme le décrivent Medley et Mitzel[24] en plus de déterminer le moment et la durée de la session d'observation, fournit à l'observateur une liste d'actions précises ou d'indices de comportements qui pourront avoir lieu ou pas, apparaître ou non, au cours de la période d'observation.

---

24. D.M. Medley et H.E. Mitzel, Measuring Classroom Behavior by Systematic Observation, in N.L. Gage (Ed.), "Handbook of Research on Teacning", Chicago, Rand McNally, 1963, p. 298

Le rapport d'observation indiquera lesquels des indices sont apparus au cours de chaque période d'observation et dans certains cas, avec quelle fréquence.

L'observateur utilisant ce protocole devra porter son attention sur tout ce qui est susceptible d'être noté selon son protocole. Cependant, le protocole ne décrivant pas nécessairement toutes les actions et ne retenant pas obligatoirement tous les indices comportementaux susceptibles d'apparaître au cours d'une période d'observation, l'observateur ne pourra tenir compte des actions et des indices qui ne font pas partie de son protocole. Il pourra passer de longs moments sans rien noter dans son rapport d'observation.

En suivant ce protocole, l'observateur est amené à travailler en suivant les cinq étapes de l'observation. A la quatrième étape, son jugement sera de l'ordre d'une alternative: le fait observé est ou n'est pas un indice accepté par le protocole.

Contrairement aux grilles et aux échelles d'observation, la durée de la session d'observation a une influence sur les résultats car ici l'observateur note chacune des apparitions d'une indice.

Le protocole d'observation d'indices répond aux besoins créés par l'observation d'une situation dynamique, d'une action pendant son déroulement.

Pour chaque indice du protocole, on obtient une fréquence. On peut comparer cette fréquence à d'autres, établir des rapports et des comportements. Ces résultats font partie de l'échelle de mesure à rapports.

La situation scolaire décrite dans la problématique renferme des situations où l'enseignant voudrait obtenir des données à propos du déroulement de certaines actions. Le protocole d'observation d'indices semble promettre une réponse à cette partie du problème. Cela justifie l'étude plus complète qui sera faite dans la prochaine section au sujet de ce procédé d'observation.

5.6 Le protocole d'observation de catégories (category system)

Comme pour la méthode d'observation précédente, le protocole d'observation de catégories établit pour l'observateur le moment et la durée de l'observation. Mais à la différence de celle-ci, tel que le précisent Medley et Mitzel[25], il permet l'enregistrement de toutes les actions qui surviennent au cours de la période d'observation.

En effet, le protocole d'observation de catégories limite l'observation à un aspect de l'enseignement, du climat de classe ou des relations maître-élève, détermine une unité d'action (en termes de temps: une minute par exemple) et définit un ensemble complet, fini, de catégories de sorte que chaque action puisse être catégorisée et ne puisse l'être que dans une seule des catégories définies.

Le rapport d'observation indiquera le nombre d'unités d'action qui ont lieu et le nombre d'unités d'actions classées dans chaque catégorie.

---

25 Idem, p. 299

L'observateur utilisant ce protocole passera par chacune des cinq étapes de l'observation. À la quatrième, il devra décider à quelle catégorie définie par le protocole appartient l'unité d'action qu'il vient d'observer. La décision ici est plus complexe, ce n'est plus une alternative. L'observateur doit chercher, par sa décision, à minimiser les erreurs de classification.

Pour obtenir les résultats, il faut établir la fréquence observée pour chaque catégorie du protocole. Ces fréquences pourront ensuite être comparées entre elles. On pourra établir des rapports et des proportions pour servir de base à l'interprétation des résultats qui, en tant que mesures, appartiennent à l'échelle de rapports.

Ce procédé d'observation est beaucoup utilisé en recherche. C'est un procédé complexe, exigeant beaucoup de précision de la part des observateurs qui doivent subir un entraînement et réussir des essais préliminaires.

Ces facteurs le rendent peu réaliste pour la situation scolaire où les enseignants font face à des groupes d'une trentaine d'élèves et où le temps qui pourrait être réservé à l'observation est limité.

5.7 Les échelles d'appréciation (rating scales)

Plusieurs termes sont utilisés pour traduire "rating scales": grille d'évaluation, terme courant chez les praticiens; échelle d'évaluation, terme proposé par Daniel Bélanger[26] dans sa traduction de Selltiz, Wrightsman et Cook; échelles d'appréciation, terme retenu par Dominique Morissette[27]. Pour les raisons exposées précédemment, c'est ce dernier terme qui sera utilisé ici.

Bien que l'observation soit à la base de l'utilisation des échelles d'appréciation, celles-ci font en plus appel au jugement et à l'expérience de l'utilisateur.

Les échelles d'appréciation fournissent à l'utilisateur une série d'actions ou de caractéristiques à observer, et une échelle sur laquelle on doit indiquer le degré de qualité ou de quantité présente ou observée par rapport à chaque action ou caractéristique de la série. Dans certains cas, ces instruments sont accompagnés de directives pour leur utilisation.

---

26. Claire Selltiz, L.S. Wrightsman, et al ..., op. cit.
27. Dominique Morissette, op. cit.

La diversité des formes d'échelles d'appréciation apparaît surtout dans la façon de présenter à l'utilisateur l'échelle des opinions, sur la qualité et sur la fréquence, que chacune des formes propose. Souvent, par ailleurs, le nom qui est donné à l'échelle renseigne assez justement sur le type d'échelle d'opinions proposé.

Avant de considérer ci-après quelques-unes des formes d'échelles d'appréciation mentionnés par plusieurs auteurs, il conviendrait d'examiner le rôle qui est assigné à l'observateur qui utilise ces échelles, puisque ce rôle est semblable d'une forme à l'autre.

5.7.1 Le rôle de l'observateur

Lorsque l'observateur utilise une échelle d'appréciation il doit suivre les cinq étapes d'une observation: attention, description, comparaison, jugement et consignation.

Après avoir été presque exclusivement observateur pendant l'étape où il doit diriger son attention sur l'objet, l'observateur se transforme graduellement en appréciateur au cours de la deuxième étape de l'observation.

Il doit apprécier, se faire une idée, une opinion à propos de ce qu'il observe. C'est cette appréciation qu'il précise et décrit avant de passer à l'étape suivante, la troisième, où il devra comparer son appréciation avec celles décrites dans l'échelle qu'il utilise. A l'étape suivante, il choisira le point de l'échelle d'appréciation qui se rapproche le plus de sa propre appréciation pour la consigner selon les règles établies, ce qui constitue la cinquième étape.

5.7.2 Le type de résultats

Ghiselli et Brown[28] établissent deux catégories générales d'échelles d'appréciation:

1) les échelles continues ou graphiques,

2) les échelles de catégories discrètes.

Dans les deux cas, les standards que l'on applique aux observations avant d'en arr ver à une appréciation que l'on pourra consigner, peuvent être les mêmes: des chiffres, des descriptions ou des adjectifs, des indications larges ou des échantillons précis.

Cette catégorisation, reprise par Gronlund[29], amène deux types de résultats.

---

28. E.E. Ghiselli et C.W. Brown, op. cit.

29. N.E. Gronlund, op. cit.

Utilisant une échelle graphique, l'observateur fait un trait sur la ligne continue se rapportant à l'élément observé. Ce trait peut être fait n'importe où sur la ligne; il ne doit pas nécessairement correspondre aux traits indiqués sur cette ligne lorsque l'échelle fournit des descriptions à titre de guide. On pourrait avoir, par exemple, le graphique suivant:

participation aux discussions

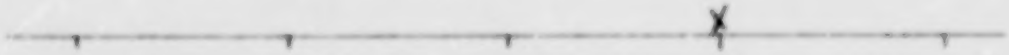
ne participe jamais, se tait, est inactif	participe autant que les autres membres du groupe	participe plus que n'importe quel autre membre du groupe

Le résultat s'obtient en plaçant une règle graduée sous la ligne et en inscrivant le nombre correspondant au trait indiqué. Dans l'exemple précédent, le résultat serait de 10,4 si l'on utilise une règle centimétrique. On peut utiliser une gradation en % ou toute autre unité arbitraire.

De cette façon, les échelles graphiques produisent des résultats continus qui se rattachent à l'échelle de mesure ordinale. Quelle que soit l'unité de mesure utilisée pour décrire la position du trait sur la ligne, le résultat obtenu devra être interprété en termes de rangs.

Par contre, lorsqu'il utilise l'échelle de catégories, l'observateur doit choisir exactement l'une ou l'autre des catégories proposées par l'échelle. On aurait par exemple:

- participation aux discussions



ne participe jamais, se tait, est inactif	participe autant que les autres membres du groupe	participe plus que n'importe quel autre membre du groupe

Pour trouver le résultat, lorsque les catégories ne sont pas numériques elles-mêmes, une série de nombre est arbitrairement assignée aux catégories. On pourrait, dans l'exemple ci-haut, utiliser l'échelle 0, 5, 10, 15, 20 et le résultat correspondant au X indiqué par l'observateur serait 15.

Ainsi, les échelles de catégories produisent des résultats discrets. Ces résultats sont aussi des mesures ordinales et devront être interprétés en termes de rangs.

Si l'on suit l'exposé de Ghiselli et Brown [30], ayant d'une part, deux types d'échelles d'appréciation:

1. les échelles graphiques (continues)
2. les échelles de catégories (discrètes)

et d'autre part quatre types de standards:

a) des chiffres ou des lettres
b) des descriptions ou des adjectifs
c) des modèles larges (échelles comparatives)
d) des échantillons précis

on pourrait retrouver huit types d'instruments, huit types d'échelles d'apprêciation.

Dans les faits, par contre, il est surtout question de trois d'entre elles:

l'échelle d'apprêciation

- graphique descriptive (1b)
- numérique ou alphabétique (2a)
- par échantillons (2d)

Après que ces échelles auront été expliquées sommairement, ce type d'instrument sera confronté aux besoins définis par la situation problématique déjà exposée.

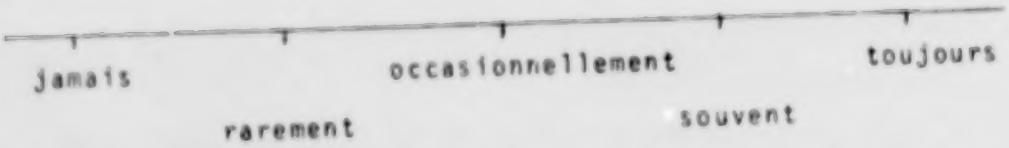
---

30. E.E. Ghiselli et C.W. Brown, op. cit.

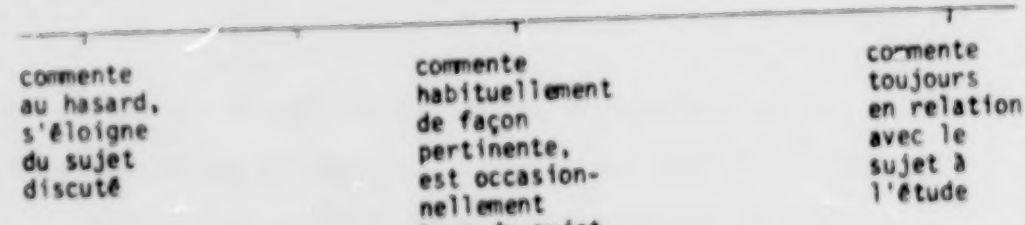
5.7.3 L'échelle graphique descriptive

L'échelle graphique présente les appréciations, à propos de la qualité ou de la fréquence observée, sous la forme d'un graphique.

Ainsi, après chaque énoncé de l'élément à observer, on peut retrouver, selon l'exemple donné par Gronlund[31]



D'autres échelles graphiques incorporent l'élément à observer avec les définitions de l'échelle, Morissette[32] lui donne alors le nom d'échelle descriptive. L'item qui suit s'inspire de ceux donnés en exemple par Gronlund (1976, p. 437).



31. Idem, p. 436

32. Dominique Morissette, op. cit., p. 190

On trouve aussi des échelles graphiques décrivant plusieurs points placés sur une ligne. Lors de l'étape comparaison, l'observateur pourrait dans ce cas comparer son opinion ou appréciation à propos de ce qu'il a observé, avec la description donnée pour chacun des points définis, avant de décider de la place de celle-ci sur l'échelle fournie par l'instrument.

5.7.4 L'échelle d'appréciation numérique ou alphabétique (alphabetical or numerical scales)

Cette échelle d'appréciation fait appel à des chiffres ou à des lettres, pour servir de bornes aux catégories de qualité ou de fréquence qu'elle propose à l'utilisateur.

Pour accompagner l'énoncé de l'élément à observer, Gronlund (1976) donne en exemple, l'échelle suivante:
1. insatisfaisant, 2. en-dessous de la moyenne, 3. moyen, 4. au-dessus de la moyenne, 5. exceptionnel.

Cet exemple illustre la difficulté de concevoir une échelle numérique simple, une échelle où chaque catégorie ne serait représentée que par un chiffre, sans plus.
On aurait alors:

- Participation aux discussions	1	2	3	4	5
- Pertinence des commentaires	1	2	3	4	5

Ce procédé est peu utilisé. Habituellement, chaque catégorie (1, 2, 3, 4, 5 ou A, B, C, D, E) est décrite par au moins un mot, de sorte que les échelles d'appréciation de ce type sont à la fois numérique (ou alphabétique) et descriptive, bien que toujours on les nomme simplement numérique (ou alphabétique) sans référence à leur partie descriptive.

Le nombre de catégories ou positions est variable d'un instrument à l'autre. Généralement, ce nombre est impair, mais Ghiselli et Brown[33] mentionnent que les observateurs consentaient plus facilement à donner leur opinion lorsque l'échelle comportait quatre catégories que lorsqu'elle n'en comportait que deux. Selon Selltiz, Wrightsman et Cook[34] "les échelles à 5 ou à 7 catégories sont celles qu'on emploie le plus fréquemment".

---

33. E.E. Ghiselli et C.W. Brown, op. cit., p. 109

34. Claire Selltiz, L.S. Wrightsman, et al ...., op. cit., p. 400

5.7.5 L'échelle d'échantillons (product scales) (behavior - sample rating scale)

Introduisant sa discussion de l'évaluation de produits, Gronlund (1976, p. 439) souligne que dans certaines domaines, tels que la dactylographie et le travail du bois, il pourrait être très souhaitable d'évaluer la méthode de travail de l'élève pendant les premières phases de l'apprentissage et d'évaluer les productions de cet élève plus tard, après que les habiletés de base aient été maîtrisées.

Pour donner une cote à ces produits, des échelles d'échantillons ont été élaborées. Elles consistent en une série d'échantillons du produit à apprécier, échantillons qui sont choisis soigneusement pour représenter différents degrés de qualité.

Pour apprécier un comportement Ghiselli et Brown[35] décrivent une méthode similaire, où les échantillons gradués de produits sont remplacés par des échantillons gradués de comportement. Pour chaque degré de l'échelle, on décrit un ou plusieurs comportements possibles.

---

35. E.E. Ghiselli et C.W. Brown, op. cit., p. 107

Pour utiliser l'échelle d'échantillons, l'observateur doit comparer ce qu'il perçoit du produit ou du comportement à apprécier avec l'échantillon de chaque degré de l'échelle et attribuer au produit ou au comportement la valeur d'échelle dont il s'approche le plus.

5.7.6 Utilisation des échelles d'appréciation

Lorsqu'elles sont utilisées en milieu scolaire, les échelles d'appréciation fournissent à l'enseignant des moyens de comparer les élèves entre eux ou de comparer leurs performances ou leurs produits à des standards définis plus globalement que ceux d'une grille ou d'une échelle d'observation.

Les échelles d'appréciation sont employées à la fin d'une activité plutôt que pendant le déroulement de celle-ci.

Les résultats qu'elles produisent ne sont pas interprétables aussi directement que ceux des grilles et des échelles d'observation. Dans ces dernières, il s'agit de présence ou d'absence des éléments recherchés, alors que les échelles d'appréciation font appel à l'opinion de l'observateur, à son appréciation d'une situation comprenant plusieurs éléments.

C'est principalement pour ces raisons que les échelles d'appréciation ne seront pas retenues comme réponse possible à la situation problématique qui appelle un instrument ou un procédé permettant, en autres, d'obtenir des informations précises sur le déroulement même d'un processus ou d'une activité.

6. Une classification des instruments

Les instruments et procédés d'observation qui viennent d'être décrits peuvent être classés à partir de chacun des deux angles sous lesquels ils ont été analysés: le rôle de l'observateur en référence aux cinq étapes de l'observation et le type de résultats produits.

Certains instruments d'observation ne font appel qu'à trois des étapes de l'observation: 1) attention, 2) description et 5) consignation. Ce type d'observation pourrait être nommée observation simple.

Par opposition, l'observation qui s'exécute en suivant la séquence complète des cinq étapes pourrait être dite structurée. La plupart des instruments et procédés étudiés relèvent de ce type d'observation.

D'autre part, l'étape jugement, la quatrième, s'accomplit de façon différente selon le type d'instrument ou de procédé avec lequel on travaille. Cela justifie une classification des instruments de l'observation structurée en trois types.

Lorsque le jugement à faire est une alternative (oui ou non, présent ou absent), l'influence de l'observateur sur le résultat est moindre que lorsque le jugement à faire est une catégorisation de faits. Et, dans le cas d'une catégorisation de faits, l'observateur a aussi un rôle moins important que lorsqu'il doit catégoriser son appréciation des faits plutôt que les faits mêmes.

Par ailleurs, les instruments de l'observation structurée peuvent être classés aussi selon le type d'information qu'ils produisent. On aurait dans ce cas deux types d'instruments: ceux qui apportent des données quantitatives et ceux qui fournissent des données qualitatives. Les termes quantitatif et qualitatif pourraient être retenus pour décrire deux classes d'instruments.

L'étape jugement de tous les instruments de type qualitatif est une catégorisation d'appréciations. Le terme qualitatif pourrait les décrire.

L'étape jugement de trois des instruments et procédés de type quantitatif est une alternative. De plus, l'ensemble des éléments qui composent les listes de ces instruments d'observation, ne recouvre pas l'ensemble des éléments qu'il serait possible d'observer au cours d'une session, pour un objet d'observation donné. L'appellation de ce type d'instruments pourrait donc comporter deux termes: celui de quantitatif, en référence au type de données recueillies, et celui de ouvert qualifiant l'ensemble des éléments observés.

Un seul procédé n'est pas inclus dans les classes définies ci-haut: le protocole d'observation de catégories. Tel que son nom le laisse entendre, ce procédé demande à l'observateur de catégoriser les faits qu'il observe en tenant compte des facteurs définis dans le protocole même. De plus, cet instrument est élaboré de façon à ce que l'ensemble des éléments observés soit un ensemble fermé, comprenant tous les éléments qu'il sera possible d'observer au cours d'une session, pour l'objet d'observation désignée. Le protocole d'observation de catégories pourrait être dit fermé par rapport à l'ensemble des éléments observés. De par le type de données qu'il produit c'est par ailleurs un instrument de type quantitatif.

La page qui suit présentera la classification proposée sous forme de tableau. A titre d'information complémentaire, l'échelle de mesure à laquelle se rattache chaque instrument sera indiquée.

Tableau 2

Une classification des instruments d'observation

Instruments d'observation simple

- l'anecdote
- le rapport d'activité

Instruments d'observation structurée

- de type quantitatif ouvert
  - la grille d'observation (o)
  - l'échelle d'observation (i)
  - le protocole d'observation d'indices (r)
- de type quantitatif fermé
  - le protocole d'observation de catégories (r)

de type qualitatif

- les échelles d'appréciation (o)
  - l'échelle graphique-descriptive
  - l'échelle numérique ou alphabétique
  - l'échelle d'échantillons

Légende pour l'échelle de mesure

o - échelle ordinale

i - échelle d'intervalles

r - échelle de rapports

CHAPITRE III

VALIDITE ET FIDELITE

Plusieurs des auteurs considérés au cours de la revue de littérature traitent des facteurs susceptibles d'avoir un effet sur les qualités de validité et de fidélité des observations. Le tableau qui suit rassemble ces facteurs autour de trois pôles: l'observateur, l'instrument et la session d'observation.

Les questions sont nombreuses et l'unanimité n'est pas faite autour des réponses à privilégier. Herbert et Attridge notent:

> "The amount of knowledge presently available that pertains to problems of observational research is extraordinarily limited".[1]

Concernant la validité et la fidélité, ces mêmes auteurs ajoutent plus loin:

> "Issues of validity, particularly the meaning and determination of reliability, are among the most contentious (litigious) in the observation litterature ...".[2]

Aucun des articles répertoriés pour la présente étude ne permet de les contredire.

---

1. J. Herbert et C. Attridge, A Guide for Developers and Users of Observation Systems and Manuals, in "American Research Journal", vol. 12, no 1, hiver 1975, p. 2

2. Idem, p. 6

Tableau 3

Les facteurs de fidélité et de validité

1. Les facteurs reliés à l'observateur:

   1.1 le degré d'entraînement (Reid, 1970; Popham 1975; Heyns et Zander, 1953; Medley et Mitzel, 1963; Boyd et VereDeVault, 1966; Travers, 1969), et le type d'entraînement (Wildman et al., 1975)

   1.2 sa justesse -- "accuracy" -- (Mash et Makohoniuk, 1975; Medley et Mitzel, 1963), basée sur sa connaissance de l'objet observé (Kerlinger, 1973)

   1.3 l'inférence (Herbert et Attridge, 1975; Kerlinger, 1973; Rosenshine, 1970)

   1.4 l'effet de halo (Medley et Mitzel, 1963)

2. Les facteurs rattachés à l'instrument:

   2.1 l'échantillonnage des éléments à observer (Heyns et Zander, 1953; Medley et Mitzel, 1963; Kerlinger, 1973), et la taille de l'unité d'observation (Heyns et Zander, 1953)

   2.2 la définition d'un cadre de référence pour l'observateur (Heyns et Zander, 1953; Boyd et VereDeVault, 1966)

3. Les facteurs rattachés à la séance d'observation:

   3.1 les conditions dans lesquelles la session se déroule (Reid, 1970), et sa longueur (Reid, 1970; Wrightstone, 1960; Medley et Mitzel, 1963)

   3.2 les effets de l'observateur sur l'observé (Medley et Mitzel, 1963; Boyd et VereDeVault, 1966; Travers, 1969; Kerlinger, 1973, Smith, 1965)

Avant de faire ici un exposé succint des questions concernant la validité et la fidélité, quelques termes sont définis afin de préciser dès maintenant le sens qu'ils prendront tout au long de l'étude.

Cet ensemble de définitions est tiré du guide sur l'observation systématique préparé par Herbert et Attridge[3], auteurs déjà cités.

Validité: degré avec lequel les mesures obtenues par un instrument décrivent réellement ce qu'elles prétendent décrire.

Fidélité: justesse (accuracy) et stabilité (consistency) des mesures obtenues par un instrument.

Observabilité: degré avec lequel les éléments à observer, inclus dans un instrument peuvent être perçus par un observateur entraîné, soit directement par ses propres sens, soit avec l'aide d'un équipement donné.

Objectivité: réfère d'une part, à la puissance de résistance d'un instrument face aux modifications qu'un observateur, selon ses préférences, attentes, besoins, sentiments et autres biais, pourrait vouloir faire,
et d'autre part, à la puissance d'influence qu'a l'instrument sur la situation ou la personne observée.

---

3. Idem

Tel que souligné, les diverses façons de garantir la validité des résultats de l'observation, fidélité incluse, ne font pas encore l'objet d'une théorie sur la méthodologie de l'observation. Ce chapitre rassemblera donc les commentaires des auteurs sur les questions figurant au tableau élaboré précédemment, en fera ressortir l'importance et présentera les solutions proposées.

1. Les facteurs reliés à l'observateur

1.1 Le degré d'entraînement

Lorsqu'on fait de l'observation, Heyns et Zander[4] soulignent qu'on utilise alors une personne comme instrument de mesure. En conséquence, le choix et l'entraînement des observateurs deviennent primordial. Ils ajoutent que les observateurs doivent:

a) être familiers avec le cadre théorique de la recherche et les objectifs de l'observation;

b) avoir vécu des expériences qui les ont sensibilisés aux dimensions recherchées par l'observation;

---

4. R.W. Heyns et A.F. Zander, Observation of Group Behavior, in L. Festinger et D. Katz (Eds) "Research Methods in the Behavioral Sciences", New-York, H.R.W., 1953, p. 416

c) avoir acquis de l'expérience avec le protocole d'observation, au cours d'essais d'entraînement préparatoires, si possible avec des groupes semblables aux groupes à venir.

Travers[5] suggère même de les faire participer au développement même du protocole d'observation alors que Popham[6] inclut, parmi les activités d'entraînement, des activités permettant de résoudre les désaccords entre les observateurs (interobserver disagreement resolution).

Le choix et l'entraînement des observateurs semblent être des questions importantes en regard de la validité des résultats de l'observation. Les suggestions proposées peuvent probablement être appliquées en milieu scolaire, malgré leur manque de précision. On peut par ailleurs supposer que l'entraînement doive durer jusqu'à qu'il y ait un degré acceptable d'accord entre les observateurs pour une situation donnée.

---

5. Robert M.W. Travers, An Introduction to Educational Research, New-York, Macmillan, 1958, p. 118

6. W. James Popham, Educational Evaluation, Englewood Cliffs, Prentice-Hall, 1975, p. 98

1.2 La justesse de l'observateur (accuracy)

Meldey et Mitzel[7] mentionnent la justesse des observations comme une des trois conditions essentielles de la validité des résultats, les deux autres étant l'échantillonnage du hasard, et le sens qui est donné aux résultats lors de leur interprétation. La justesse des observations repose en bonne partie sur la justesse de l'observateur et Kerlinger[8] note à cet effet que celui-ci doit connaître l'objet qu'il observe et le connaître de façon compétente.

Ce commentaire de Kerlinger rejoint l'idée de prévisibilité du comportement à observer que relèvent Mash et Makohoniuk[9] en même temps que les notions de fréquence et complexité des catégories utilisées par l'instrument pour décrire ce comportement.

Ces facteurs, tout en n'étant pas directement du ressort de l'observateur, influencent la justesse des

---

7. D.M. Medley et H.E. Mitzel, Measuring Classroom Behavior by Systematic Observation, in N.L. Gage (Ed.), "Handbook of Research on Teaching", Chicago, Rand McNally, 1963, p. 245

8. F.N. Kerlinger, Foundations of Behavioral Research, New York, H.R.W., 2e éd., 1973, p. 534

9. Eric J. Mash et George Makohonuik, The Effects of Prior Information and Behavioral Predictability on Observer Accuracy, in "Child Development", vol. 46, no 2, 1975, p. 513

observations qu'il fait. Ainsi, lorsque l'observateur connaît bien les manifestations possibles d'un comportement, celles-ci deviennent prévisibles pour lui. Ce comportement est observable dans le sens défini au début du chapitre. Par contre, s'il s'agit d'un comportement difficilement observable, pour lequel on ne connaît pas bien les différentes manifestations possibles, l'observateur ne pourra les prévoir, certaines passeront même inaperçues. Les observations faites dans un tel cas seront beaucoup moins fidèles, beaucoup moins justes; la justesse de l'observateur sera aussi moins grande.

Lorsque le comportement à observer apparaît très fréquemment, il sera difficile pour l'observateur d'en relever toutes les manifestations avec justesse. Si les comportements à observer sont de nature complexe et que, par conséquent, les catégories pour les classer sont complexes également, l'observateur pourra éprouver des difficultés à reconnaître et à classer ces comportements.

Ces facteurs de complexité des catégories, de fréquence, de complexité et de prévisibilité du comportement à observer se traduisent, pour l'observateur, en deux types de problèmes: l'inférence et l'effet de halo.

1.3 L'inférence

Selon la définition donnée par Rosenshine[10], l'inférence c'est le processus qui intervient entre les données objectives vues ou entendues et la codification de ces données sur un instrument d'observation. Kerlinger[11], pour sa part, note que la validité des résultats est en relation directe avec la marge d'interprétation laissée à l'observateur par l'instrument ou le protocole d'observation.

Rosenshine, comparant le protocole d'observation de catégories avec les échelles d'appréciation, conclut que le protocole laisse, en général, peu de marge d'interprétation à l'observateur:

> "Category systems are classified as low-inference measures because the items focus upon specific, denotable, relatively objective behaviors ... and because these events are recorded as frequency counts. Rating systems are classified as high-inference measures because they lack such specifity."[12]

Il semble donc que le degré d'inférence dépende beaucoup plus de l'instrument utilisé que de l'observateur.

---

10. B. Rosenshine, Evaluation of Classroom Instruction, in "Review of Educational Research", vol. 40, 1970, p. 281

11. F.N. Kerlinger, op. cit., p. 539

12. B. Rosenshine, loc. cit., p. 281

Celui-ci cependant, devrait être avisé qu'en ne suivant pas les consignes données, il risque d'augmenter ce degré d'inférence, ce qui aurait pour conséquence d'augmenter d'autant l'erreur de mesure.

1.4 L'effet de halo

L'autre problème avec lequel l'observateur est confronté est celui de l'effet de halo qui est la tendance d'une personne à être influencée par une autre caractéristique que celle qu'il veut observer ou par son impression générale de l'objet ou de la personne sous observation.

L'effet de halo se rencontre surtout dans les situations d'appréciation. Mais il peut jouer dans les situations d'observation également. Medley et Mitzel[13] étudient cette possibilité. Ils constatent que lorsque les items d'un instrument d'observation sont construits de façon à laisser place à l'interprétation, l'effet de halo influencera les résultats des observations en transformant ceux-ci en des opinions ou réactions de l'observateur. Ils concluent: "The solution of the problem clearly depends on the construction of items."[14]

---

13. D.M. Medley et H.E. Mitzel, loc. cit., pp. 305-306

14. Idem

Ici encore, on peut dire que l'observateur ne sera pas le premier responsable. Son travail consiste à consigner des observations et non des impressions ou des appréciations. La qualité de l'instrument qu'il utilisera semble donc de première importance.

De même que pour la questi n d'inférence, il serait sounaitable d'informer l'observateur de ce qu'est l'effet de halo et de lui indiquer les conséquences de celui-ci sur la validité des résultats des observations.

2. Les facteurs reliés à l'instrument

Dans le prochain chapitre qui présentera de façon détaillée les trois types d'instruments d'observation retenus, les moyens de rédiger les items pour minimiser les possibilités d'inférence et les effets de halo seront décrits. Il sera question ici des autres facteurs de fidélité et de validité qui sont rattachés à l'instrument.

2.1 L'échantillonnage et l'unité d'observation

L'échantillonnage des éléments à observer de même que la taille de l'unité d'observation sont mentionnés par plusieurs auteurs comme facteurs essentiels de la validité d'un instrument ou d'un protocole d'observation.

Heyns et Zander[15] considèrent qu'il faut spécifier clairement la taille de l'unité d'observation afin que tous les observateurs utilisant le même instrument le fassent dans le même sens. De cette façon l'accord entre observateurs (interobserver agreement) sera maintenu. Généralement, la taille de l'unité d'observation est spécifiée en même temps que l'échantillonnage des éléments d'observation.

Pour effectuer l'échantillonnage, on se base soit sur l'événement ou l'acte, soit sur le temps. Wrightstone[16] décrit l'échantillonnage par le temps

> "Time sampling consists of observation of behavior of an individual or a group for definite short periods of time and the recording of occurrence or nonoccurrence of objectively defined forms of behavior during each period."[16]

.l rapporte à titre d'exemple une étude sur les contacts sociaux d'enfants d'âge pré-scolaire. Le protocole d'observation spécifiait d'observer un enfant pendant cinq minutes, puis de passer au suivant, et ainsi de suite. L'ordre dans lequel un enfant était observé, était prédéterminé et modifié

---

15. R.W. Heyns et A.F. Zander, loc. cit.

16. J.W. Wrightstone, Observational Techniques, in C.W. Harris et M.R. Liva (Eds), "Encyclopedia of Educational Research", 3e ed., New-York, Macmillan, 1960, p. 928

chaque jour de façon à ce que l'on ait pour chaque enfant un dossier complet, soit, des données pour un premier cinq minutes, un deuxième cinq minutes, etc.

Boyd et Vere DeVault rapportent que Wright (1959)[17] utilisa un échantillon d'intervalles de quinze secondes: le comportement d'un premier intervenant était observé pour quinze secondes puis l'observateur prenait les quinze secondes suivantes pour consigner son observation. Il passait ensuite à l'observation d'une autre intervention.

Ces mêmes auteurs mentionnent la technique d'échantionnage par le temps utilisée par Bonney (1955)[18] qui a utilisé vingt échantillons de cinq minutes, choisis au hasard, pour chaque enfant.

Boyd et Vere DeVault rapportent enfin que Flanders (1965) décrit sa technique d'échantillonnage comme suit:

> "At the end of each 3 second period, he (observer) decides which of a prescribed of numbered categories best represents the communication events just completed."[19]

---

17. R.D. Boyd et M. Vere DeVault, The Observation and Recording of Behavior, vol. XXXVI, no 5, décembre 1966, p. 542

18. Idem

19. Idem

Ainsi donc, ce type d'échantillonnage impose une contrainte de temps quant au moment où l'observateur doit faire une observation. Ces contraintes peuvent être d'application systématique de même qu'aléatoire.

Comparant cette technique avec celle de l'échantiollonnage par événements, Selltiz, Wrightsman et al., concluent que la première a l'avantage de fournir un échantillon plus représentatif du comportement, qu'elle serait valide pour les comportements qui surviennent fréquemment. Par contre, les échantillons qui en résultent n'ont pas "le caractère de continuité et d'authenticité apparente de l'échantillonnage par événement."[20]

Ce dernier commentaire semble rejoindre ce que pensaient Medley et Mitzel[21] lorsqu'ils nommèrent unités naturelles (natural units) les unités d'observation obtenues avec l'échantillonnage par événement.

C'est ce type d'échantillonnage qu'utilise le protocole d'observation d'indices comme le soulignent Medley et Mitzel[22], cette technique peut aussi servir dans

---

20. Claire Selltiz, L.S. Wrightsman et al., Les méthodes de recherche en sciences sociales, Montréal, H.R.W., 1977, p. 281 (Traduit par David Bélanger)

21. D.M. Medley et H.E. Mitzel, loc. cit. p. 301

22. Idem

le protocole d'observation de catégories. Les données obtenues indiquent quels événements ont eu lieu et, dans certains cas, combien de fois. Un événement c'est une phrase, un contact, une question, lorsqu'il s'agit d'actions ou de comportements.

Quelques-unes des façons de procéder à l'échantillonnage d'événements décrites par Altmann (1973) citée par Selltiz, Wrightsman et al.[23], pourraient s'avérer utiles à la présente recherche:

- Toutes les fois qu'un comportement donné se présente au cours de la période d'échantillonnage, on enregistre le fait. On y arrive en concentrant son attention soit sur un individu ou sur un petit groupe d'individus.

- Echantillonnage un-zéro. Noter la présence ou l'absence (et non la fréquence) d'un comportement au cours de la période d'échantillonnage. Ces périodes sont souvent courtes (dix à quinze secondes) et sont consécutives. Il est important de noter que l'échantillonnage un-zéro donne les fréquences des intervalles où le comportement se produit et non pas la fréquence du comportement lui-même, ni le nombre de fois où il se présente durant chaque intervalle.

---

23. Claire Selltiz, L.S. Wrightsman et al., op. cit., pp. 282-283

- Echantillonnage instantané. Il s'agit de porter son attention sur un individu seulement assez longtemps pour observer et enregistrer un comportement puis de passer à l'observation de la personne suivante. Le grand avantage: observer un grand nombre d'individus dans de courtes périodes de temps. Cet échantillon peut donner le pourcentage de temps consacré à une activité quelconque: diviser le nombre des échantillons où cette activité s'est présentée par le nombre total des échantillons dans la période d'observation. L'interprétation de ce pourcentage est plus directe que celle de l'échantillonnage un-zéro.

Medley et Mitzel[24] notent que l'échantillonnage un-zéro permet d'observer des comportements très fréquents, pour lesquels on ne peut enregistrer toutes les apparitions. Ces auteurs ont utilisé la technique un-zéro et ils concluent qu'il semble que le nombre de fois qu'un indice apparaît au cours d'une période de quelques heures soit en forte corrélation avec le nombre de courtes périodes de temps où ce même indice apparaît au moins une fois.

L'échantillonnage des éléments à observer a été réalisé de plusieurs façons différentes. Quatre exemples

---

24. D.M. Medley et H.E. Mitzel, loc. cit., p. 302

d'échantillons-temps et trois exemples d'échantillons-événements ont été présentés. Quant à la décision de procéder à un échantillonnage par le temps ou par événements, il semble que celle-ci doive reposer sur une étude de la situation à observer. Cette décision doit tenter d'assurer aux observations le plus de ressemblance possible avec l'objet observé en regard de sa continuité, sa fréquence, sa nature même.

2.2 Le cadre de référence

Un autre facteur relevé comme très important pour la validité des résultats de l'observation réside dans la définition d'un cadre de référence pour les observateurs. Plusieurs caractéristiques d'un instrument ou d'un protocole d'observation sont incluses dans ce terme.

Heyns et Zander[25] mentionnent que trois décisions sont prises par ceux qui ont élaborés l'instrument ou le protocole:

a. jusqu'à quel point l'observateur doit-il tenir compte du contexte qui entoure l'action ou le comportement observé?

b. l'observateur doit-il tenir compte des intentions

---

25. R.W. Heyns et A.F. Zander , loc. cit., p. 416

de la personne observée lorsqu'il classifiera ses observations au sujet de cette personne?

c. jusqu'à quel point les conséquences des actions observées doivent-elles influencer leur classification par l'observateur?

Boyd et DeVault[26] expriment cette dernière question d'une autre manière. Ils considèrent qu'un comportement ou une action comporte deux aspects: une action est l'expression d'une intention de même que son instrument. Ils supposent que chaque aspect peut être observé et enregistré séparément et suggèrent que l'utilisateur d'un instrument d'observation décide d'obtenir des informations indistinctement sur les deux aspects ou alors sur l'un ou l'autre de manière précise. Autant de questions auxquelles le cadre de référence doit apporter une réponse.

Outre un exposé clair à propos du point de vue qu'il doit adopter et conserver, le cadre de référence de l'observateur doit aussi comporter des indications suffisantes au sujet du cadre théorique sur lequel l'instrument ou le protocole s'appuie, lorsqu'un tel cadre théorique existe. Dans le cas contraire, la démarche de la recherche doit être

---

26. R.D. Boyd et M. Vere DeVault, loc. cit., p. 534

expliquée à l'observateur pour qu'il puisse situer son rôle dans celle-ci.

Le cadre de référence oriente les observateurs, il les guide tous dans le même sens en leur faisant comprendre ce que l'on attend d'eux. Il les aide aussi à mieux connaître les éléments qu'ils auront à observer, connaissance qui est essentielle pour en faire l'observation. Pour bien accomplir ses fonctions auprès de l'observateur, on peut supposer que le cadre de référence doit être le plus clair et le plus complet possible.

3. Les facteurs reliés à la session d'observation

3.1. Les conditions de l'observation

Les conditions dans lesquelles se déroule la session d'observation, telles que l'endroit, l'horaire et la longueur de la session, font l'objet d'une simple énumération de la part de Reid[27].

Wrightstone[28] pour sa part indique que la durée d'une observation peut se situer entre quelques minutes et plusieurs heures. Les facteurs qui déterminent cette durée,

---

27. John B. Reid, Realiability Assessment of Observational Date: A Possible Methodological Problem, in Child "Development", vol. 41, no 4, 1970, p. 1149

28. J.W. Wrightstone, loc. cit., p. 928

serait la nature du phénomène étudié et la practicabilité (feasability). Il y a aussi les limites imposées par la division des activités scolaires en périodes de classe. Cet auteur suggère que pour observer le comportement de jeunes enfants, spécialement en situation libre, on organise plusieurs courtes périodes d'observation (de cinq à quinze minutes), avec rotation dans l'horaire de la journée et de la semaine.

Medley et Mitzel[29] soutiennent pour leur part qu'une visite dans une classe du primaire dure généralement entre vingt et trente minutes, et que dans une classe du secondaire il semble logique qu'une visite coïncide avec la durée totale d'une période de cours, soit entre quarante et soixante minutes.

Le seul critère qui a été relevé est donc celui de l'adaptation. Les conditions de l'observation doivent être adaptées à la nature de l'objet observé de même qu'aux conditions d'ordre pratique dans lesquelles se déroule l'observation. On peut supposer que les conditions d'utilisation d'un instrument ou d'un protocole d'observation seront soumises à un groupe d'utilisateurs probables, pour être critiquées et améliorées si besoin est, avant d'être fixées de façon définitive.

---

29. D.M. Medley et H.E. Mitzel, loc. cit., p. 304

3.2 Les effets observateur-observé

Une question fort controversée est celle des effets de l'observateur ou de l'observation sur la personne observée.

Medley et Mitzel[30] étudient la relation entre l'observateur et l'observé pour conclure qu'il est possible de prendre au moins deux précautions pour éviter que les comportements de la personne observée ne soient pas représentatifs de ceux qu'elle aurait si elle n'était pas sous observation. Ils suggèrent d'abord qu'il n'y ait pas plus d'un ou deux observateurs présents en classe. Ils ajoutent que les personnes impliquées dans l'observation (la personne observée, son supérieur de même que l'observateur) soient informées et comprennent les motifs pour lesquels les données sont recueillies et la façon dont elles seront utilisées.

Boyd et Vere DeVault[31] étudient aussi la relation entre observateur et observé et concluent pour leur part que l'influence peut être réciproque. Celle que la personne observée peut avoir sur l'observateur pourrait se traduire, on l'a vu, par un effet de halo sur les résultats.

---

30. Idem, p. 307

31. R.D. Boyd et M. Vere DeVault, loc. cit., p. 544

Les effets de l'observation ou de l'observateur sur l'observé sont appelés réactivité. Ils sont catégorisés par Webb, Campbell et al.[32]:

1. L'effet cochon d'Inde -- les personnes qui se rendent compte qu'on les prend pour sujets d'expérience ne se comportent pas selon leur façon habituelle. Elles montrent le meilleur d'elles-mêmes ou alors elles ont une réaction hostile.

2. Le choix d'un rôle -- chaque personne a un répertoire de comportements qui lui permet de jouer un certain nombre de rôles différents. Le rôle qu'elle choisit dépend des indices sociaux présents dans la situation.

3. Les apprentissages suscités par la situation de recherche -- La rencontre d'un chercheur ou la nouveauté de la situation peut susciter des apprentissages au cours des premières sessions d'observation et de ce fait rendre invalides les observations subséquentes.

4. Les tendances (response set) -- La plus connue est celle du répondant qui acquiesce sans cesse avec l'intervieweur; une autre consiste à agir ou répondre de la façon la plus socialement désirable.

Pour minimiser ces effets, les auteurs suggèrent

---

32. E. Webb, D.T. Campbel, et al., Unobtrusive Measures, Chicago, Rand McNally, 1966, 225 p.

de remplacer l'observation directe par des observations dissimulées (unobtrusive), ou indirectes.

A titre d'exemple d'observation dissimulée, ils mentionnent, entre autres, les traces laissées par l'usure, le passage ou l'utilisation, ou encore les dossiers tenus par les institutions, hôpitaux, écoles, ou autres.

Dans le même ordre d'idée, Brandt[33] indique que l'observation des produits et des outils, en l'absence de la personne qui a réalisé ces produits et utilisé ces outils, est un moyen de mesure dissimulée (unobtrusive).

Kerlinger[34] conclut dans le même sens, affirmant que les effets de l'observation sur l'observé peuvent être réels. Cependant, si l'observateur prend soin de ne pas s'imposer et évite de donner aux personnes observées l'impression d'être jugées, alors l'observateur en tant que stimulus est presque nul.

Les conclusions des travaux de Smith, McPhail et al.,[35] rejoignent celle de Kerlinger. Leur étude de la littérature

---

33. R.M. Brandt, Studying Behavior in Natural Settings, New York, H.R.W., 1972, p. 201

34. F.N. Kerlinger, op. cit., p. 538

35. R.L. Smith, C. McPhail, et al., Reactivity to Systematic Observation with Film: a Field Experiment, in "Sociometry", vol. 38, no 4, 1975, pp. 536-550

n'a révélé aucune preuve pour appuyer ou rejeter définitevement l'hypothèse de réactivité. Celle-ci, selon eux, aurait trois sources: les activités de l'observé, les activités de l'observateur et la possibilité pour l'observé de sentir la présence d'un observateur.

Plus l'observé est concentré sur sa tâche, et plus cette tâche fait appel à tous ses sens, moins celui-ci sera influencé par l'observateur.

Les activités de l'observateur et les instruments qu'il utilise doivent être jugés le moins incongrus possible par les observés pour minimiser la réactivité.

Enfin, la proximité de l'observateur est une source certaine de réactivité. Autant que faire se peut, l'observateur devrait être suffisamment éloigné de l'observé pour que celui-ci ne puisse le voir, l'entendre écrire ou l'entendre bouger, et ne puisse sentir sa présence près de lui.

Il semble donc possible de réduire les effets probables de l'observation ou de l'observateur sur la personne observée. Tous les auteurs qui ont relevé cette question, suggèrent en effet des précautions à prendre pour pallier à la situation créée par l'observation. La principale d'entre elles, pour l'observation directe, semble être que l'observateur fasse sentir sa présence le moins possible.

4. Les indices de fidélité

Dans les pages qui précèdent, un ensemble de moyens pour garantir la fidélité des résultats de l'observation a été présenté. On peut supposer que ces moyens s'appuient d'abord sur un consensus d'experts pour juger de la façon dont ils sont exploités, et pour juger de leur à-propos, de leur suffisance et de leur efficacité.

Il faut rappeler cependant que dans le cas de l'inférence et de l'effet de halo, certains auteurs ont proposé quelques techniques précises. Celles-ci seront étudiées dans le chapitre suivant.

En plus du recours à un consensus d'experts, certains indices mathématiques ont été développés pour indiquer le degré probable de fidélité des résultats de l'observation. Selon Herbert et Attridge, il s'agit bien ici de la fidélité des résultats de l'observation et non de la fidélité de l'instrument, ou du protocole, ou des observateurs.

> "The position we adopt is the traditional one, namely, that reliability, in the measurement-theoretic sense, is a property of measures obtained through the application of a system, not a property of an instrument, nor of a system, nor of a record, nor of observers, though qualities of each of these constrain the realiability of measures obtained."[36]

---

36. J. Herbert et C. Attridge, loc. cit., p. 14

Les indices de fidélité relevés dans la littérature ne sont pas des estimations de la fidélité des résultats de l'observation. On tente généralement d'établir la fidélité des observateurs par le biais d'un coefficient d'accord (observer agreement). On tente aussi d'obtenir des informations sur la stabilité des comportements observés. Peu d'auteurs proposent une méthode pour estimer la fidélité des résultats mêmes de l'observation.

Ces trois différents indices de fidélité sont étudiés ci-après.

4.1 Le coefficient d'accord

Ce premier indice est défini par Medley et Mitzel[37] comme étant la corrélation entre des scores basés sur des observations faites par différents observateurs, au même moment.

Flanders[38], qui définit la fidélité de son instrument uniquement en termes de coefficient d'accord entre les

---

37. D.M. Medley et H.E. Mitzel, loc. cit., p. 253

38. N.A. Flanders, The Problems of Observer Training and Reliability, in E. Amidan et J. Hough (eds), "Interaction Abalysis: Theory and Application", Mass., Addison- Wesley, 1967, cité par B. McIaw, et al., loc. cit., p. 15

observateurs, calcule celui-ci différemment. Utilisant le coefficient de Scott (1955), il parvient à estimer jusqu'à quel point le degré d'accord entre les observateurs dépasse celui qu'il aurait obtenu par hasard simplement.

Quelle que soit la façon de calculer le coefficient d'accord entre les observateurs, Medley et Mitzel critiquent fermement l'utilisation de cet indice de fidélité. Il leur semble impropre pour diverses raisons, d'interpréter le coefficient d'accord en termes de fidélité des résultats.

Ainsi, cette dernière peut être faible même dans le cas où le coefficient d'accord est fort. Medley et Mitzel illustrent deux occasions où une telle situation se retrouve:

> "... observers might be able to agree perfectly on the number of seats in a room, yet if the number of seats in all rooms is equal, or nearly so, the reliability of seat counts as a measure of differences between classes will be zero."39

Dans cette situation, on a un accord parfait entre les observateurs mais le trop peu de variété entre les objets observés (le nombre de sièges par classe) produit un coefficient de fidélité nul.

---

39. Idem, p. 310

> "Near-perfect agreement could also be reached about the number of boys in a room wearing red neckties; but if every boy changed the color of his necktie every day, the reliability of these counts would be zero."[40]

Cette fois-ci encore, on a un accord presque parfait entre les observateurs. Dans cette situation, c'est la trop grande variabilité entre les objets observés qui réduit à zero le coefficient de fidélité des résultats.

Il sembre donc impossible d'utiliser le coefficient d'accord entre observateurs comme indice du degré de fidélité des résultats de l'observation. Cette opinion est reprise par McGaw, Wardrop, et al.[41] ainsi que par Herbert et Attridge[42]. Ces derniers notent toutefois, que, malheureusement les chercheurs utilisent encore le coefficient d'accord entre observateurs comme indice de fidélité des résultats.

4.2 Le coefficient de stabilité

Un deuxième indice mathématique est le coefficient de stabilité dont Medley et Mitzel rapportent la définition

---

40. Idem, p.310

41. V.McGaw, J.L. Wardrop, et al., Classroom Observation Schemes: Where are the Errors?, in "American Educational Research Journal", 1972, p. 14

42. J. Herbert et C. Attridge, loc. cit., p. 14

admise couramment: "A correlation between scores made on observations made by the same observer at different times ... tells us something about the consistency of behavior from time to time."[43]

Le coefficient de stabilité suppose un plan d'observation dans lequel un observateur va, à plusieurs moments différents, observer une personne. Entre l'un et l'autre moment, aucun apprentissage organisé n'aura eu lieu pour la personne observée: les différences de comportement de cette personne sont donc supposées aléatoires.

Egalement, d'une session d'observation à une autre, on suppose aléatoires les changements de la situation dans laquelle se déroule l'observation. Il pourra s'agir d'une matière d'enseignement différente ou encore d'un type d'activités différent.

Si l'on accepte la définition de stabilité donnée ci-haut, on interprète la variabilité d'un comportement dans le temps comme étant une erreur aléatoire de mesure, erreur que l'on peut attribuer soit à la situation au cours de laquelle l'observation a eu lieu, soit à la personne

---

43. D.M. Medley et H.E. Mitzel, loc. cit., p. 254

observée, soit à ces deux facteurs de variance.

Il est toutefois possible de penser qu'un comportement puisse varier vraiment, pour s'adapter à une situation. C'est sur cette idée que McGaw, Wardrop et Brenda s'appuient pour critiquer la position qui vient d'être décrite.

> "In the case of observations of teacher behavior efforts develop indices to characterize particular teachers appears to be misplaced unless there is some allowance made for lawful adaptations of behavior to different situations."[44]

La définition d'un coefficient de stabilité ne devrait porter, selon ces auteurs, que sur des comportements manifestés à différentes occasions, à l'intérieur d'un même type de situation.

> "Variability among situations, within teacher, would no longer be attributed to error variance but to true variance, i.e., variation in the measure reflecting real, systematic variation in the person."[45]

Il semble donc que le coefficient de stabilité ne rende compte que d'une partie de la variance des résultats de l'observation mais qu'il ne puisse suffire pour informer sur la fidélité des résultats obtenus.

---

44. B. McGaw, J.L. Wardrop, et al., loc. cit., p. 16

45. Idem, p. 17

Sur ce point, tous les auteurs étudiés se rejoignent Medley et Mitzel l'expliquent clairement lorsqu'ils écrivent: "... a stability coefficient does not estimate the accuracy of a score ... since it is based on a correlation between observations made by a single observer."[46]

4.3 Le coefficient de fidélité

Dans la théorie classique des tests, on ne calcule qu'un seul coefficient de fidélité des résultats pour estimer le degré de corrélation qu'il y aurait entre deux séries de résultats provenant de deux administrations d'un même test.

La variabilité des résultats de l'observation pouvant être influencée, on l'a vu, par de multiples facteurs, la littérature revue suggère qu'on calcule plusieurs indices de fidélité.

La fidélité est définie par Herbert et Attridge de la façon suivante:

> "Reliability refers to the accuracy and consistency with which the measures obtained by an instrument actually describe what they purport to describe ... the same results from similar events though these events are separated in time or location, or have different participants or settings."[47]

---

46. D.M. Medley et H.E. Mitzel, loc. cit., p. 254

47. J. Herbert et C. Attridge, loc. cit., p. 6

Deux groupes d'auteurs étudient de façon assez complète la question de l'estimation de la fidélité des résultats de l'observation et proposent des solutions précises.

Medley et Mitzel[48], en 1963, à la suite de leurs recherches, concluent que l'analyse de la variance d'un ensemble de données (obtenues par deux observateurs, en deux occasions différentes, au sujet d'un même objet d'observation) donne une estimation non-biaisée de la fidélité des résultats analysés. De plus, ils affirment que cette estimation est plus précise que n'importe quelle combinaison de corrélations intra-classes qui seraient calculées avec ces mêmes données. Enfin, une seule analyse de la variance permet d'estimer tous les différents coefficients de fidélité jugés appropriés aux divers usages que l'on voudra faire des résultats d'une observation.

Le design général qu'ils proposent pour l'estimation de la fidélité est une analyse de la variance à quatre facteurs de classification, ces facteurs étant (1) le nombre de classes observées, (2) le nombre d'observateurs, (3) le nombre d'items et (4) le nombre d'occasions ou de sessions

---

48. D.M. Medley et H.E. Mitzel, loc. cit., p. 310

d'observation. Une telle analyse de la variance produit des carrés moyens qui sont utilisés comme coefficients de fidélité.

Le deuxième groupe d'auteurs reprennent la question quelque dix ans plus tard. McGaw, Wardrop et Brenda, en 1972 concluent:

> "To attempt to assess each of the sources of error variance in a complex model such as this ... the most appropriate approach would be to use a variance components analysis such as that proposed by Medley and Mitzel (1963) or by Cronbach and his associates ..."[49]

Ces auteurs appuient donc la démarche exposée ci-haut. Par ailleurs, ils étudient la théorie de la généralisabilité développée par Cronbach et ses associés et en démontre l'application possible au problème d'estimation de la fidélité des résultats de l'observation.

5. La validité

La fidélité des résultats de l'observation étant un des premiers facteurs de leur validité, on peut d'abord mentionner que le fait de tenir compte de l'ensemble des conseils qui viennent d'être relevés, pour assurer la fidélité, c'est déjà garantir une certaine part de la validité.

---

49. B. McGaw , J.L. Wardrop, et al., loc. cit., p. 17

> "A measure is valid to the extent that differences in scores yielded by it reflect actual differences in behavior ..."[50]

Il est généralement admis qu'on ne puisse établir la validité d'un résultat que s'il est d'abord fidèle à la réalité qu'il tente de quantifier.

Quelques auteurs traitent spécifiquement de validité dans leur étude de l'observation. Kerlinger suggère que la validité prédictive des résultats de l'observation soit établie:

> "A simple aspect of the validity of observation measures is their predictive power. Do they predict any relevant criteria dependably? The trouble as usual is in the criteria."[51]

Un deuxième type de validité lui semble important:

> "The important clue to the study of validity of behavioral observation measures would seem to be construct validity."[52]

la validité de construit est également mentionnée par Herbert et Attridge qui déplorent le peu d'informations disponibles concernant la validité:

> "System designers and users frequently do not go far enough in the development of their tools to establish validity ..."[53]

---

50. D.M. Medley et H.E. Mitzel, loc. cit., p. 250

51. F.N. Kerlinger, op. cit., p. 539

52. Idem, p. 540

53. J. Herbert et C. Attridge, loc. cit., p. 15

Selon ces auteurs, différents types de validité pourraient être établis pour les résultats de l'observation. En plus de la validité de construit déjà mentionnée, ils signalent qu'il serait bon d'établir, selon les possibilités, la validité apparente (face validity), la validité de contenu et la validité concourante (concurrent criterio-related validity).

D'autre part, Herbert et Attridge, rappelant que la validité apparente est rattachée au degré par lequel un instrument paraît mesurer ou décrire ce qu'il prétend mesurer ou décrire, concluent:

> "It is determined through the application of the intelligence, intuition, and reason of the developer, objective colleagues, and users and, at the stage we are presently in our validation of observation techniques, it is an important characteristic for an instrument to possess."54

Il semble donc important de tenter d'établir au moins la validité apparente des résultats obtenus par observation. Quant aux moyens d'y arriver, de même qu'aux façons de démontrer les autres types de validité relevés ci-haut, les auteurs consultés pour la présente recherche n'en décrivent aucun qui soit spécifique à la situation de

---

54. Idem, p. 16

l'observation. Cela pourrait confirmer l'opinion de Rosenshine et Furst: "Most of the instruments developed in the observation field have yet to prove the validity of their measures."[55] Les théoriciens de même que les créateurs des instruments d'observation, semblent avoir omis la question de la validité des résultats.

L'observation systématique est un procédé complexe. Minimiser l'erreur de mesure signifie tenter de contrôler de multiples facteurs possibles de variance, facteurs qui influencent le degré de fidélité des résultats. De plus, démontrer la validité de ces mêmes résultats c'est chercher à rassembler le plus d'informations possible au sujet de leur statut par rapport à des résultats comparables ou à des critères, par rapport à une théorie ou un ensemble d'hypothèses de recherche.

Pour unifier ces multiples questions, pour mieux définir les solutions possibles, on peut penser qu'une théorie de l'observation sera développée au cours des années à venir.

---

55. B. Rosenshine et N. Furst, The use of direct observation to study teaching, in R.M.W. Travers (Ed.), "Second Handbook of Research on Teaching", Chicago, Rand MaNally, 1973, p. 125

CHAPITRE IV

TROIS INSTRUMENTS ET QUELQUES EXEMPLES

Des neuf instruments introduits précédemment, trois ont été retenus parce que susceptibles d'apporter des solutions à l'intérieur de la problématique précisée au départ. Ce sont: la grille d'observation, l'échelle d'observation et le protocole d'observation d'indices.

Chacun d'eux a été étudié sommairement dans le cadre du deuxième chapitre et tous se retrouvent dans la même catégorie, selon la classification qui a été proposée. Ce sont des instruments d'observation structurée, de type quantitatif ouvert. Ce dernier qualificatif a été donné aux instruments qui décrivent certaines des manifestations de certains comportements et non pas toutes les manifestations possibles de tous les comportements pour une session d'observation donnée.

L'analyse plus complète entreprise maintenant vise à faire ressortir un ensemble de critères dont il faudra tenir compte au moment de proposer une méthodologie pour leur utilisation en classe. Toutes les informations recueillies au sujet de ces trois instruments seront exposées et analysées en ce sens.

1. Les grilles d'observation

Le plus simple des instruments d'observation structurée, la grille d'observation, est étudié par plusieurs auteurs. Les éléments de leur étude se résument à la présentation de la grille et de quatre ou cinq directives pour sa construction, ainsi qu'à des commentaires sur les avantages et les inconvénients de son utilisation.

La grille d'observation ayant déjà été présentée à la section 5.3 du deuxième chapitre, il ne sera question ici que des techniques pour sa construction.

Trois auteurs donnent des directives pour la construction de la grille d'observation: Morissette[1], Tenbrink[2] et Gronlund[3]. L'unanimité est faite autour de quatre points, décrits par Gronlund.

"1. Identify and describe clearly each of the specific actions desired in the performance.

---

1. Dominique Morissette, Les examens de rendement scolaire, Québec, P.U.L., 1979, 390 p

2. T.D. Tenbrink, Evaluation: A Practical Guide for Teachers, New-York, McGraw-Hill, 1974, 493 p.

3. N.E. Gronlund, Measurement and Evaluation in Teaching, 3e ed., New-York, Macmillan, 1976, 590 p

2. Add to the list those actions which represent common errors, if they are limited in number and can be clearly identified.
3. Arrange the desired actions and likely errors in the approximate order in which they are expected to occur.
4. Provide a simple procedure for numbering the actions in sequence or for checking each action as it occurs".[4]

On demande d'abord une description claire du comportement ou de l'action ou de la caractéristique à observer. Le degré de clarté de cette description est en rapport inversement proportionnel avec le degré d'inférence de l'observateur. C'est donc là un moyen d'assurer la fidélité des observations.

Inclure dans la liste des éléments à observer des erreurs communes peut être utile tant pour noter une performance que pour recueillir des informations pour intervenir par l'enseignement correctif auprès d'un élève pour améliorer sa performance. La qualité de celle-ci peut être diminuée à la fois par des omissions et par des erreurs, d'où l'intérêt d'inclure les deux dans la grille.

---

4. Idem, pp. 446-447

Ordonner les éléments à observer selon leur ordre probable d'apparition facilitera la tâche de l'observateur. La grille le distraira moins de sa tâche principale qui est d'observer avec le plus de justesse possible. En ce sens, la disposition des éléments contribuera à la fidélité des observations.

La quatrième directive prévoit deux façons de noter les observations. Lorsque la séquence des actions posées est importante, l'observateur reçoit la directive d'écrire à côté de chaque action le rang auquel elle est apparue. Dans le cas contraire, l'observateur est requis d'indiquer un chevron, un x ou autre marque simple, à côté de l'action observée.

Aucune étude de la grille d'observation ne mentionne de précautions à prendre ou de directives à suivre en regard de l'entraînement des observateurs, de l'échantillonnage, du cadre de référence, des conditions de l'observation et des effets observateur-observé. Si cela était, la fidélité et la validité des observations pourraient probablement avoir de meilleures garanties.

Enfin, aucune procédure de quantification des résultats n'est suggérée par les auteurs recensés au cours de cette recherche.

D'après cette étude, les principaux critères d'observation sont ceux de clarté des items et de simplicité de la procédure. Pour assurer la validité et la fidélité des résultats, il pourrait être possible de tenir compte des conclusions des différentes sections du chapitre précédent, lors de l'élaboration et de l'utilisation des grilles d'observation.

2. Les échelles d'observation

Introduites à la section 5.4 du deuxième chapitre, les échelles d'observation présentent, comme les grilles, une liste d'éléments à observer. Ce que l'échelle ajoute, c'est la pondération des éléments.

Cette technique a été présentée par Ghiselli et Brown en 1955 dans le cadre d'un manuel de psychologie spécialisée (Personnel and Industrial Psychology). Proposée pour être utilisée avec le personnel employé dans les industries, cette procédure n'a pas été reprise par la psychologie de l'éducation, ce qui explique probablement que les sources de la présente étude de l'échelle d'observation soient limitées aux deux auteurs précités.

La grille d'observation est utilisée dans le domaine de l'éducation. La similarité de l'échelle d'observation avec celle-ci pourrait justifier son étude ici. De plus, la particularité de l'échelle d'observation, la pondération, semble rendre cette technique intéressante comme réponse possible au problème posé au début.

L'élaboration d'une échelle d'observation peut se diviser, selon Ghiselli et Brown, en six étapes consécutives.

> "1. The job is analyzed in order to discover the major areas in which success should be measured. Some knowledge about the importance of each of these areas also is obtained. This step is necessary, first of all, as the basis for the collection of items, and secondary, as a means of checking the thoroughness of coverage of the job by the final scale".[5]

L'analyse de la tâche est faite dans le but de déterminer les éléments qu'il faudra observer. En plus d'en retirer une certaine connaissance de l'importance relative de ces éléments, le constructeur, en suivant de près les données de cette analyse, assurera une part de la validité de contenu de son instrument.

---

5. E.E. Ghiselli et C.W. Brown, Personnel and Industrial Psychology, New-York, McGraw-Hill, 1955, p. 110

> "2. A large number of statements are collected concerning the performance of individuals on the job".[6]

Ghiselli et Brown suggèrent de collectionner un grand nombre d'items concernant l'objet à observer. Parallèlement à leur proposition d'interview des contremaîtres, dans le domaine scolaire la collection pourrait se faire auprès des enseignants et des rédacteurs des programmes. La terminologie utilisée par les enseignants pourra faciliter l'étape qui suit.

> "3. The statements are edited to ensure that are specific and to the point. They should not be double-barreled in meaning but should be so formulated that they can erther be accepted or rejected".[7]

Les items conservés pour être soumis aux juges, devront être rédigés dans des termes spécifiques et précis, sans ambiguïté, et formulés de façon à ce que l'observateur y réagisse par oui ou non, présent ou absent. De tels items permettent de prévoir et d'observer le comportement ce qui contribue à la justesse de l'observateur, facteur qui contribue à la fidélité des observations. De même, c'est à cette étape que sera déterminée le degré d'inférence qui sera laissé à l'observateur.

---

6. Idem, p. 111

7. Idem

> "4. A number of judges are asked to evaluate each of the edited statements on a scale relative to the degree of job success a worker so described would have. ... The scale value of each item is obtained by taking the average of all judges' ratings. In using the final scale, the scale values are not reproduced on the form and are unknown to the rater."[8]

Chaque juge évalue tous les items qui lui sont soumis. Il exprime son avis, sur une échelle de cinq à onze points, indiquant de cette façon le degré de contribution de chaque élément au succès de la performance. En calculant la moyenne des valeurs accordées par les juges, on obtient, pour chaque item une valeur d'échelle. Ces valeurs ne sont pas inscrites sur le formulaire qu'utilise l'observateur dans le but de ne pas nuire à son objectivité.

> "5. Those items on which the judges show wide disagreement in their ratings are discarded. Obviously they are ambiguous and have different meanings for the various judges. Only items are retained on which there is a high degree of agreement".[9]

> "6. As a final step, statements are selected that represent all levels of ability on the scale and which also cover all aspects of the job as indicated by the job analysis".[10]

---

9. Idem, p. 112

10. Idem

Ces deux dernières étapes de l'élaboration concernant la mise en forme de l'échelle d'observation. On retrouve ici encore le souci de conserver des items qui ne sont pas ambigus de même que des items qui représentent tous les différents niveaux d'habileté. On travaille ainsi, de façon systématique, à la fidélité des observations, à leur validité de contenu et à leur validité apparente.

Tel que déjà spécifié, pour utiliser l'échelle d'observation, il faut cocher tous les éléments dont on observe la présence, toutes les caractéristiques qui décrivent l'objet observé. Le résultat est la moyenne des valeurs des éléments marqués par l'observateur.

Selon Ghiselli et Brown, l'instrument élaboré par cette technique peut être utilisé pour évaluer la performance d'un employé. Aucune précision n'est faite sur les conditions dans lesquelles cette évaluation doit se dérouler.

Des études de fidélité de certaines échelles d'observation, développées pour des performances spécifiques, démontrent des résultats que Ghiselli et Brown qualifient de surprenants. Ces études de fidélité reposent sur des données obtenues par deux observateurs indépendants. Il s'agit d'un coefficient d'accord dont on a montré, plus avant, la faiblesse en tant qu'indice de fidélité.

Etant donné ses qualités de construction, l'échelle d'observation pourrait être un instrument plus fidèle que la grille d'observation. Cette technique pourrait permettre de travailler à obtenir un consensus sur l'échelonnage d'un ensemble de manifestations d'une habileté par rapport à différents niveaux de maîtrise de cette habileté. Ce consensus établi par un groupe d'enseignants, leur permettrait d'établir un cheminement possible pour développer cette habileté, d'observer et de situer périodiquement leurs élèves en regard de la maîtrise de cette habileté, pour les aider dans leur cheminement.

Il faudrait toutefois tenter de démontrer la fidélité des résultats de même que leur validité concourante, à défaut de pouvoir assurer leur validité de construit. Il faudrait également répondre aux questions se référant à l'entraînement de l'observateur, à l'échantillonage des éléments à observer et aux conditions de l'observation pour assurer le plus de fidélité possible aux observations faites par le biais de l'échelle d'observation.

Les principaux critères identifiés dans l'étude qui précède concernent l'élaboration de l'échelle. Les items doivent être formulés de façon précise, sans ambiguïté, en termes tels que l'observateur pourra répondre oui ou non.

Parmi les items ayant retenu l'accord des juges, on en choisira pour représenter le mieux possible l'ensemble des différents niveaux de performance. Un critère est à retenir pour l'utilisation: les valeurs des items ne sont pas indiquées sur le formulaire que l'observateur aura en main.

3. Le protocole d'observation d'indices

Le protocole d'observation d'indices (sign system), déjà introduit à la section 5.5 du deuxième chapitre, est une technique d'observation utilisée en recherche scientifique.

Des auteurs recensés pour la présente étude, deux décrivent le protocole d'observation d'indices de façon exhaustive: Medley et Mitzel. Ils notent que le premier problème à résoudre est celui de décider quels comportements et quels aspects des comportements il faudra noter et concluent à ce sujet: "The purpose for which a study is undertaken usually determines which behaviors are to be observed by indicating them in greater or less detail".[11] Ils ajoutent que l'observation d'indices est souvent utilisée dans des situations que la recherche théorique n'a pas encore définies.

---

11. D.M. Medley et H.E. Mitzel, Measuring Classroom Behavior by Systematic Observation, in N.L. Gage (Ed.), "Handbook of Research on Teaching", Chicago, Rand McNally, 1963, p. 298

Il faut noter que l'ensemble des indices recherchés n'est pas la totalité des manifestations possibles d'un comportement ou d'une caractéristique. C'est ce qui a permis, plus tôt, de qualifier cette technique d'ouverte.

Lorsqu'on a amassé un certain nombre d'idées sur ce qu'on désire observer, il faut transformer ces idées en items ou indices. Des indications très précises sur la formulation de ce type d'items sont données par Medley et Mitzel.

D'abord, l'indice doit être significatif par lui-même. Un protocole d'observation d'indices peut rassembler au-delà de soixante-dix indices. Il est impensable de songer à accompagner chaque indice d'un court paragraphe pour le définir. La justesse de l'observateur en serait diminuée.

> "If, when he sees a behavior, he has to stop to remember a definition ... he may well miss the next behavior that he is supposed to record."[12]

La formulation même des indices doit présenter trois qualités principales: 1) le verbe d'action doit être au présent pour que cette action puisse être consignée au moment où elle se produit 2) l'action doit être positive donc observable

---

12. Idem, p. 302

pendant la session d'observation -- une omission ne peut être constatée qu'à la fin de la session 3) le complément d'objet de l'action observée doit être au singulier -- s'il s'agissait d'un pluriel, l'observateur devrait attendre que l'action se produise plus d'une fois avant de la consigner, et décider lui-même de la signification de ce pluriel: deux, trois, ...

L'arrangement des indices sur le formulaire qui sera utilisé par l'observateur doit avant tout faciliter la consignation des observations; en deuxième lieu, la disposition des items doit viser à minimiser l'effet de halo. Viennent ensuite des critères concernant la facilité de calculer le résultat et de se former une impression générale de celui-ci par un examen rapide du formulaire.

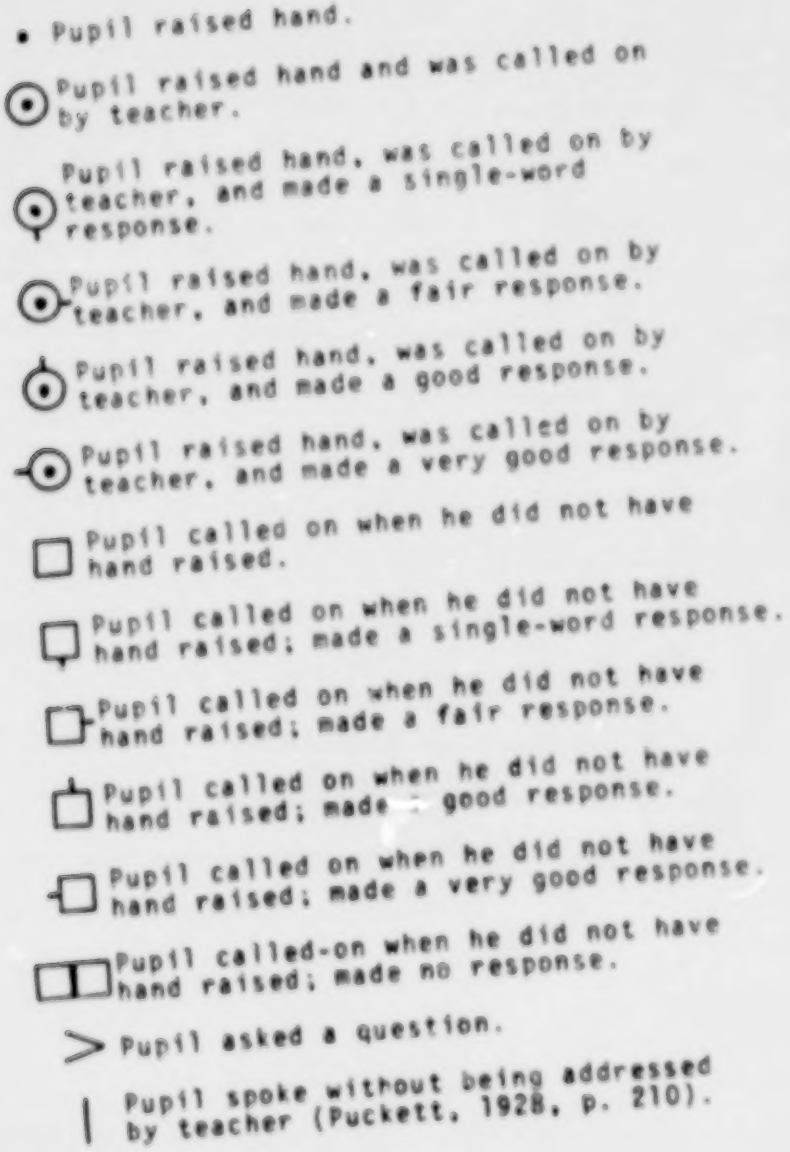
La consigne de Medley et Mitzel sur une disposition des indices visant à diminuer l'effet de halo, démontre une préoccupation pour la fidélité des observations. C'est une consigne très large cependant; elle n'est pas accompagnée d'exemples qui permettraient de comprendre le type de disposition souhaitée.

Medley et Mitzel précisent par ailleurs qu'il est préférable que la consignation des observations se fasse par le biais d'un même signe (un crochet, une croix, etc.) à côté des items qui forment une liste, plutôt que par un ensemble de signes différents qu'on disposerait sur une même ligne. "It seems preferable, whenever possible, to identify different behaviors by changing the position rather than the form of the tallies".[13]

Les tableaux qui suivent illustrent ces deux types de présentation des indices. Le tableau 4 présente un ensemble de signes différents utilisés par R.C. Puckett en 1928. Selon les consignes relevées ici, cet exemple n'est pas à retenir. D'abord parce que les indices sont rédigés au passé composé. Ensuite parce que malgré son apparente simplicité ce système exigera de l'observateur beaucoup plus de mémoire et de précision au moment de la consignation de ses observations. Le rythme et la justesse de l'observateur pourrait être influencé par ce mode de consignation.

Tableau 4

L'ensemble des signes d'un protocole[a]



a. R.C. Puckett, cité par D.M. Medley et H.E. Mitzel, loc. cit., p. 254

Le tableau 5 montre une partie d'un protocole d'observation d'indices qui demande à l'observateur de tracer un x à l'endroit approprié.

TABLEAU 5

(Sns)	I	III	V	Tot
S2 t mvs frly				
S3 p mvs frly				
S5 t cls p dr, etc				
S6 t shws afct f p				
S7 p shws afct f t				
S8 p shws ho f t				
S9 p shws ho t p				
S10 uses srcsm				
S11 t yls				
Check				

Ce protocole prévoit six sessions d'observation, trois par observateur. C'est une portion du formulaire du premier observateur qui est représentée ici; les indices sont numérotés et inscrits en abrégé.

---

a. D.M. Medley et H.E. Mitzel, loc. cit., p. 279

Les recherches de Medley et Mitzel ont montré qu'il est préférable que le texte du formulaire soit peu compréhensible pour un non-initié: "A sign should be printed 'in toto' on the schedule, with abbreviations used to disguise the exact content".[14]

Quelques remarques sur l'utilisation d'un protocole d'observation de signes sont faites par les auteurs. Ainsi, lorsqu'on emploie un tel protocole, il arrive de longs moments ou l'observateur n'ait rien à consigner. Au sujet du comportement observé, ils indiquent: "The numer of different aspects of behavior that can be sampled at one time with this type of item is very large".[15] Ces deux remarques sont importantes à cause de leurs conséquences possibles sur l'utilisation de ce protocole en situation scolaire. Si l'observateur n'est pas constamment occupé par l'observation d'une personne, on peut supposer qu'il pourra observer plus d'une personne à la fois. Il est intéressant de noter aussi la possibilité d'observer un grand nombre d'indices à la fois.

---

14. Idem, p. 305

15. Idem, p. 299

Lorsqu'on utilise le protocole d'observation en situation de recherche il faut retenir, selon ces deux auteurs, deux considérations lors de la détermination du design:

> "(1) It is desirable that every observer record in every classroom on each section the uses; and
>
> (2) it is desirable to have two recorders use each section of the schedule, eventhough this means that each recorder must learn to use two sections, to lessen the effects of bias".[16]

Il faut se rappeler que ces considérations sont faites dans le cadre d'une analyse des résultats par l'anlayse de la variance avec quatre facteurs de classification. Les chercheurs qui voudront utiliser la théorie de la généralisabilité tel que proposé par McGaw, Wardrop et Bunda[17] pour analyser les résultats pourront probablement modifier le design de l'observation proposé ici, pour l'adapter à l'univers de résultats pour lequel ils voudront généraliser les résultats observés.

---

16. Idem, p. 304

17. B. McGaw, J.L. Wardrop, et al., Classroom Observation Schemes: Where are the Errors?, in "American Educational Research Journal", vol. 9, no 1, 1972, p. 49

Dans leur présentation du protocole d'observation d'indices, Medley et Mitzel discutent des questions concernant l'entraînement des observateurs, l'effet de halo, l'échantillonage, les effets de l'observateur sur l'observé et les conditions de la session d'observation. Les réponses qu'ils proposent à ces questions ont été rapportées précédemment, dans le cadre du chapitre qui traite de validité et fidélité. Il ne semble pas utile de les reprendre ici. Quant à la validité des résultats, ces auteurs rappellent que pour être valides, les résultats doivent d'abord être fidèles et annoncent que c'est cet aspect de la validité qu'ils traitent dans leur article. Ils mentionnent aussi la nécessité d'établir la validité concourante mais ne rapportent pas d'essais en ce sens.

Pour l'élaboration du protocole d'observation d'indices, on peut retenir, des pages qui précèdent, quelques critères:

- l'indice doit être significatif par lui-même,
- le verbe d'action doit être au présent,
- l'action observée doit être positive,
- le complément d'objet de l'action doit être au singulier,

- sur le formulaire les indices doivent être écrits en abrégé,
- chaque indice occupe une ligne,
- les indices doivent être disposés pour faciliter la consignation et éviter l'effet de halo.

4. Quelques exemples d'utilisation

Pour la présente recherche, il serait intéressant d'étudier quelques exemples d'observation d'élèves, en petits groupes. Peu d'expérimentations de ce type sont rapportées dans la littérature; quatre d'entre elles, qui se rapprochent le plus de la situation problématique, seront étudiées dans les pages qui suivent.

Le but principal de cette étude est de vérifier si l'observation structurée et systématique est praticable dans la situation d'une classe régulière. La question des qualités métrologiques ne sera pas discutée ici, puisque les critères concernés ont été rassemblés au chapitre précédent.

4.1 Learning Accomplishment Profile (LAP)

Développé par le Bureau de l'éducation des enfants handicapés de Washington, l'objectif du Learning Accomplishment Profile[18] est de remettre à l'enseignant de l'enfant handicapé un dossier simple, faisant référence à des critères, décrivant les habiletés actuelles de cet enfant.

L'instrument est divisé en trois parties:

a) un dossier des habiletés pour six facettes du développement: motricité, motricité fine, habiletés sociales, auto-suffisance, habiletés intellectuelles et langage;

b) un guide d'hiérarchisation du développement des habiletés;

c) quarante-quatre unités d'apprentissage.

Seule la première partie de cet instrument ainsi que les procédures qui s'y rapportent seront présentées ici.

La première partie du LAP comporte 42 pages. Les items ne sont pas numérotés mais on peut en estimer le nombre à près de 450. Le tableau qui suit reproduit une page de cet instrument qui s'apparente à la grille d'observation.

---

18. Etat de Washington, Bureau of Education for the Handicapped, Learning Accomplishment Profile (LAP), 1976, 135 p.

TABLEAU 6

Extrait du Learning Accomplisment Profile

FINE MOTOR (cont.)

Source	Behavior	Age (Dev.)	Assessment Date		Date of Achievement	Comments (Criteria, materials, Problems, etc.)
3	Manipulates egg beater	27 mos.	+	9/18/73		Whipped soap suds Teacher held handle
9	Enjoys finger painting	30-35 mos.	+	9/20/73		Finger pointed on formica. table - 10 min.
9	Makes mud and sand pies	30-35 mos.	+	9/24/73		Made sand pies using tea set
13	Paints strokes, dots, and circular shapes on easel	30-35 mos.	+	9/27/73		Imitated teacher w/½" brush
6	Cuts with scissors	35 mos.	+	10/ 2/73		Cut ½" partially cut strips.
13	Picks up pins, thread, etc., with each eye separately covered	36-48 mos.	-	10/ 3/73		
7	Drives nails and pegs	36-48 mos.	-	10/ 4/73		These will become
13	Builds tower of nine cubes	36-48 mos.	-	10/ 5/73		objectives for this
7	Holds crayon with fingers	36-48 mos.	+	10/ 8/73		child
3	Strings 4 beads	36-48 mos.	-	10/ 8/73		
13	Can close fist and wiggle thumb in imitation, R & L	36-39 mos.				
11	Puts 6 round pegs in round holes on pegboard	36-48 mos.				

NOTE: The child has demonstrated a developmental age of 35 months in fine motor skills. Failure on 4 of 5 items represents ceiling.

Mark + for positive demonstration of skill
Mark - for negative demonstration of skill

Les items du profil proviennent d'une revue de littérature. Une bibliographie est remise à l'utilisateur de l'instrument. La source bibliographique de chaque item étant indiquée en première colonne, l'enseignant s'y référant peut trouver des informations et commentaires utiles.

Les consignes pour l'utilisateur sont d'indiquer + lorsque l'habileté a été démontrée par l'enfant et - dans le cas contraire. L'enseignant inscrit également la date d'évaluation et ajoute des commentaires s'il y a lieu.

Les items ont été disposés sur le formulaire selon un ordre croissant de difficulté. Le niveau de performance de l'élève est déterminé par le fait d'obtenir quatre fois la mention "moins" dans une suite de cinq items. Dans ce cas, l'enseignant arrête l'observation de cette facette du développement pour l'enfant concerné.

Le manuel de l'utilisateur suggère à l'enseigant des situations d'observation:

> "Skill area: Fine Motor.
>
> Behavior: Strings four beaks (36-48 months).
>
> Activity: Give each child four beads of at least drawn on it. Have the child match the beads to the same ones on the card, and then string them on a lace.

Behavior: Imitates folding and creasing paper three times (48-60 months).
Activity: Make paper hats by folding paper rectangles".[19]

Une dizaine de situations de ce genre sont décrites. L'enseignant procède par observation informelle au cours des deux ou trois premières semaines de l'année. Ce faisant, il doit se familiariser avec l'ensemble des items du LAP: "In addition to designing specific evaluative activities, familiarity with the items on the LAP facilitates general observation skills which enchance informal assesment".[20] On compte ainsi qu'avec le temps, lorsqu'il connaîtra bien son instrument, l'enseignant augmentera son habileté à observer de façon informelle.

Cette expérience montre une façon d'utiliser l'observation pour vérifier et guider le développement d'habiletés chez de jeunes handicapés. Les documents publiés sur ce programme ne rapportent pas de façon exacte, toutes les conditions de l'utilisation de l'observation: Les enseignants bénéficient-ils de sessions d'information et d'entraînement?

---

19. Etat de Washington, Bureau of Education for the Handicapped, Anne R. Sanford (Ed.), A Manual for Use of the Learning Accomplishment Profile, 1976, p. 11

20. Idem, p. 12

Combien ont-ils d'élèves dans leur classe? Combien y a-t-il d'heures de classe par semaine?

L'absence de telles informations laisse peu d'éléments de cette expérience qui permettent de vérifier les possibilités pratiques d'utiliser l'observation en classe, dans une situation régulière ou l'enseignant est seul avec une trentaine d'élèves.

4.2 Cleveland Perceptual-Motor Rating Scale

Dans le cadre d'un projet de développement des habiletés perceptivo-motrices touchant les écoles publiques de Cleveland, un instrument du type échelle d'appréciation a été développé: le Cleveland Perceptual-Motor Rating Scale.[21]

Cet instrument a été utilisé à deux fins: obtenir des informations sur chaque enfant entrant en première année pour déterminer lesquels d'entre eux ont besoin d'aide spécifique, et, à la fin du programme d'aide, vérifier le succès de ce programme.

---

21. R. Garvey et T.R. Kimball, Development of a Perceptual-Motor Rating Scale for Classroom Teachers to Record Behavior, conférence donnée en mars 1978, 62e rencontre annuelle, American Educational Research Association, 35 p.

L'échelle comporte 28 items. L'enseignant doit faire un x dans une des trois cases dessinées en regard de chaque item:

- OK si l'action décrite ne pose pas de problème à l'enfant;
- ?? si l'action décrite pose quelquefois mais rarement, des problèmes à l'enfant observé;
- PR lorsque l'action s'avère un réel problème pour l'enfant.

L'enseignant doit noter tous les enfants de son groupe, sur chaque item, au cours d'environ vingt jours de classe. Il y a une période d'évaluation en octobre et en mai.

A titre d'illustration, voici quelques items de l'instrument.

"05. Holds pencil correctly.
09. Pastes correctly.
10. Buttons, unbuttons owr clothing.
13. Ties -- shoe laces, hats, etc.".[22]

---

22. Idem

Il a été mentionné plus avant que l'observation est à la base de l'utilisation d'une échelle d'appréciation. Aussi, quoique reposant sur une échelle d'appréciation, instrument qui n'a pas été retenu pour répondre à la situation problématique de départ, l'expérience, décrite ci-haut illustre l'applicabilité de l'observation dans le contexte d'une classe régulière.

Cependant, comme pour l'exemple précédent, toutes les conditions de l'utilisation de l'instrument ne sont pas indiquées dans le rapport de Garvey et Kimball. Il est toutefois intéressant pour cette recherche-ci de constater que l'expérience s'est déroulée dans des classes régulières et que ce sont les enseignants de ces classes qui ont complété les formulaires. Par contre, une question importante pour la présente recherche reste sans réponse: les enseignants ont-ils complété le formulaire de l'échelle d'évaluation en présence des enfants évalués ou l'ont-ils fait de mémoire, peu de temps après avoir observé un ou quelques élèves? Ce sont là des façons de procéder qui influencent directement l'organisation de la classe pour la durée de l'opération.

4.3 Maryland Systematic Teacher Observation Instrument

Développé pour appliquer la loi de l'Etat du Maryland sur l'identification hâtive des incapacités pouvant mener à des troubles d'apprentissage, le Maryland Systematic Teacher Observation Instrument[23] est utilisé pour chaque enfant s'inscrivant en première année de l'école primaire.

Cet instrument établit un sommaire de la performance de l'élève et peut être complété par l'enseignant sans modifier le déroulement des activités régulières de la classe. Les résultats sont traités par informatique et sont retournés aux écoles. Le principal est le premier responsable de l'exploitation des résultats. Il est appuyé par les enseignants et une équipe de gestion des ressources éducatives (educational management team) présidée par le principal. Les professionnels au service de l'école font partie de cette équipe. Des exemples de devis de programme d'enseignement complètement l'ensemble. L'école fait appel aux parents lorsqu'un élève semble éprouver des difficultés.

---

23. Etat du Maryland, Montgomery County Public Schools, Early Identification of Learning Disabilities and Suggestions to Assist Classroom Teachers in Program Design, 1976, 72 p.

Le rapport de cette expérience comporte des informations sur plusieurs aspects: procédures pour l'interprétation des scores, proposition de stratégies d'enseignement, renseignements sur l'élaboration de l'instrument, etc. A titre de vérification de l'applicabilité de l'observation en classe régulière, seuls l'instrument même et son utilisation par l'enseignant seront présentés ici.

Le Maryland Systematic Teacher Observation Instrument compte un total de trente-six items divisés en cinq secteurs de développement: psychomotricité, sensations et perceptions, langage, connaissances, affectivité et motivation.

Au cours des huit premières semaines de classe, l'enseignant observe ses élèves. On lui fournit un formulaire "Class Worksheet" sur lequel il peut prendre note de ses observations. Ce formulaire comporte trente-sept colonnes. A l'exception de la première, réservée aux noms des élèves, chaque colonne a pour titre un des trente-six items de l'instrument. L'enseignant peut y consigner des informations pour un groupe d'au plus trente-cinq élèves.

A la fin de la période d'observation, l'enseignant remplit un formulaire individuel pour chacun de ses élèves, jugeant de la fréquence du comportement décrit par l'item.

A titre d'exemple, voici quelques items:

	always	often	sometimes	seldom	never
"17. Gives own name and age when asked.	○	○	○	○	○
18. Stares into space.	○	○	○	○	○
19. Can identify colors (red, yellow, blue, green) by name".[24]	○	○	○	○	○

On donne à l'enseignant les consignes suivantes:

- il n'est pas nécessaire de "tester" chaque élève individuellement,
- il n'est pas nécessaire que tous les enseignants utilisent les mêmes mises-en-situation,
- si l'enseignant se pose des questions au sujet d'un enfant, il lui faudra peut-être l'observer particulièrement.

En regard de l'échelle de fréquence, on donne des définitions à l'utilisateur, tout en lui disant: "approach the task from a common sense point of view"[25], remarque qui est soulignée dans le texte. Le continuum de fréquence est le suivant:

" ... a teacher will assign:

---

24. Idem, p. 3

25. Idem, p. 2

ALWAYS: to behavior which occurs overwhelmingly often but not nessarily every single time

OFTEN: to behavior which occurs more times than not

SOMETIMES: to behavior which falls in the middle of the continuum indicating that there is a 50-50 chance that the behavior will occur

SELDOM: to behavior which doesn't usually occur but can be expected to appear now and then

NEVER: to behavior which occurs so rarely that it is almost nonexistent."[26]

A cause du type de décision que l'utilisateur doit faire, cet instrument, selon les définitions données au deuxième chapitre, est plus une échelle d'appréciation de la fréquence de comportements qu'une simple grille d'appréciation.

La procédure d'utilisation marque cependant une différence entre la période où l'enseignant observe et celle où il transcrit son appréciation sur le formulaire individuel.

---

26. Idem

Par ailleurs, le formulaire de groupe prévoit des espaces pour jusqu'à trente-cinq élèves, et c'est sur ce système d'observation que repose l'application d'une loi de l'Etat du Maryland.

Cela laisse supposer qu'une telle opération d'observation - appréciation par les enseignants de classes régulières est possible et qu'elle fournit des informations importantes pour tenter de prévenir les difficultés d'apprentissage.

Cependant, comme pour les deux illustrations qui précèdent, le rapport de cette expérience ne fait aucune mention de l'attention qui aurait été donnée à l'entraînement des observateurs, à la vérification de la justesse de leurs observations, ainsi qu'à la réduction de l'effet de halo. Des commentaires sur ces trois éléments auraient permis une étude plus complète de l'applicabilité dans nos écoles, des procédures utilisées par les écoles publiques du Maryland.

4.4 Le protocole d'Anderson

Un dernier exemple, la mise en situation faite par Anderson, sera étudié ici, à cause des possibilités qu'elle offre comme réponse au problème déterminé au départ.

La recherche d'Anderson [27] avait pour sujet les comportements des élèves, comportements reliés aux objectifs pédagogiques de l'enseignant. Il observait les comportements verbaux et non verbaux des élèves. Pour être consigné, un comportement devait être relié à une activité dont l'élève était l'initiateur (self-directed activity - SDA), activité qui se rapportait aux objectifs de l'enseignant, sans être directement influencée par celui-ci. Le protocole d'observation d'indices a permis à ce chercheur de déterminer jusqu'à quel point chacun des élèves s'adonnaient à cette activité autonome. Selon le protocole qu'il a établi, un observateur recueillait les données pour cinq élèves simultanément.

Le tableau 7 reproduit une fiche de consignation utilisée lors de cette recherche.

---

27. W.W. Anderson, A Study of the Effects of Self-directed Activity upon Quantity, Quality, and Variety of Responses in a Group Directed Reading-Thinking Activity, thèse de doctorat, non publiée, University of Virginia, 1971, cité par R.M. Brandt, "Studying Behavior in Natural Settings", New-York, Holt, 1972, pp. 102-104

Tableau 7

Reproduction d'une fiche de consignation[a]

a. R.M. Brandt, Studying Behavior in Natural Settings, New-York, Holt, 1972, p. 103

L'observateur devait suivre les consignes données. Les quatre indices recherchés étaient des comportements autonomes de l'élève observé, verbaux (V) ou non verbaux (N), directement relié aux objectifs pédagogiques (V_2 ou N_2) ou indirectement reliés à ceux-ci (V_1 ou N_1). "Asking a child for a ruler needed to do an arithmetic problem would be coded, V_1, while looking up an item of information in a reference book would probably be classified as N_2".[28]

Lorsqu'un élève n'avait aucun des quatre types de comportements recherchés, l'observateur indiquait un tiret dans l'espace prévu. La période d'observation proprement dite durait six minutes et quarante secondes. Les premières minutes passées en classe étaient consacrées à l'identification de la fiche de consignation et à la mémorisation de la position des élèves observés dans le diagramme de la classe. L'observateur devait noter ses observations à toutes les dix secondes.

On peut supposer qu'à chaque période de dix secondes, l'observateur portait son attention sur l'élève A, notait mentalement l'indice observé, portait son attention sur l'élève B, notait mentalement l'indice, et ainsi de suite.

---

28. R.M. Brandt, Studying Behavior in Natural Settings, New-York, Holt, 1972, p. 104

Il consignait ensuite, dans l'ordre, les indices relevés. Pour la période 2:50 à 3:00 de la fiche de consignation reproduite plus avant, il aurait noté mentalement à 2:51 - N_1 ; à 2:52 - rien; à 2:53 - rien; à 2:54 - rien; à 2:55 - N_1 ; ce qui a donné sur sa fiche A - N_1 ; B --; C --; D --; E - N_1. Il aurait pu prendre jusqu'à cinq secondes pour consigner ses observations. On peut supposer ici que les observateurs ont bénéficié d'une session d'entraînement.

Brandt fait rapport du coefficient de stabilité du comportement observé et du coefficient d'accord entre les observateurs. Ces données sont mentionnées ici simplement à titre indicatif.

> "This coding procedure illustrates an attempt on several pupils simultaneously. Stability coefficients from one day to the next were 0.61 for groups and 0.50 for individuals, whereas coefficients of observer agreement were above 0.90 for both individual and group measurement".[29]

Cette expérience illustre pour la présente recherche, la possibilité d'obtenir, par le biais de l'observation, des données sur plusieurs élèves à la fois. Il faut noter cependant que c'est la seule expérience du genre relevée dans la littérature. Ce type d'échantillonnage pourrait convenir à la situation problématique définie.

---

29. Idem

Il faudrait cependant que les activités se déroulant en classe au moment de l'observation, libèrent l'enseignant de son rôle d'animateur pour lui permettre d'assumer celui d'observateur.

Des quatre exemples réels étudiés dans le présent chapitre, aucun ne fournit donc un modèle qu'il serait possible de reproduire tel quel pour observer le développement des habiletés d'une trentaine d'élèves. Chacun illustre cependant des possibilités qui pourraient s'intégrer à la méthodologie de l'observation en classe qui sera proposée au chapitre suivant.

CHAPITRE V

PROPOSITION D'UNE STRATÉGIE D'OBSERVATION SYSTÉMATIQUE DES ÉLÈVES EN CLASSE

Au cours des chapitres précédents, des données concernant l'observation ont été rassemblées et analysées. Quatre essais faits dans des situations se rapprochant quelque peu de la situation du groupe-classe, ou confirmant une caractéristique particulièrement intéressante pour la situation scolaire, ont été étudiés. De plus, trois instruments ou procédés d'observation offrant des possibilités d'utilisation en classe, ont été décrits de façon détaillée. Ces analyses et études visaient à faire ressortir un ensemble de critères d'ordre théorique, critères auxquels devraient répondre un instrument ou un procédé d'observation, pour tenter d'obtenir une certaine fidélité et validité des résultats obtenus.

De façon habituelle, la littérature définit des critères en fonction d'une situation de recherche, où la seule tâche de l'observateur est celle d'observer. Certaines suggestions des auteurs semblaient cependant ouvrir des avenues pour faciliter l'utilisation de l'observation dans le cadre de l'enseignement à un groupe d'élèves. Pour concrétiser l'objectif de la présente recherche, ces avenues ont été soulignées au passage dans les pages qui précèdent.

Le but de ce chapitre est de rassembler ces suggestions et de proposer une stratégie de l'observation en classe qui répondrait aux critères identifiés en fonction de la fiabilité des résultats de cette observation.

La proposition exposée ci-après repose de plus sur une connaissance du milieu scolaire. Elle tentera donc de tenir compte à la fois des données théoriques étudiées dans les chapitres précédents et de données provenant d'un vécu en classe. Il n'en demeure pas moins que plusieurs éléments de cette proposition sont présentés à titre d'hypothèses puisque la stratégie proposée n'a pas fait l'objet d'expérimentations.

Les instruments d'observation, retenus pour répondre à la problématique soit, la grille d'observation, l'échelle d'observation et le protocole d'observation d'indices, produisent différents types de résultats. Ce fait influencera l'interprétation que l'on fera des résultats et constituera un facteur à considérer lors du choix de l'instrument. Cependant, les trois instruments présentés ont en commun une caractéristique fondamentale: chacun a pour base une liste: on ne prend pas note de tout ce que l'élève fait, on note seulement la présence de qualités, de caractéristiques ou d'actions recherchées.

Les techniques et procédures pour élaborer la liste de ce qu'on pourrait nommer les indices, sont différentes d'un instrument à l'autre. Elles sont relativement simples pour la grille d'observation, tandis que l'élaboration du protocole d'observation d'indices est plus complexe. Selon l'instrument d'observation qu'il aura choisi, l'observateur aura donc en main une liste simple ou une liste plus complexe, mais ce sera toujours une liste d'indices ou de points dont il lui faudra noter la présence ou l'absence.

On pourrait donc utiliser, tenant compte des réserves exprimées plus haut, l'un ou l'autre des trois instruments considérés afin d'observer un même objet. C'est pourquoi la stratégie proposée ici sera générale; elle pourra être appliquée soit avec la grille d'observation, soit avec l'échelle d'observation, soit enfin avec le protocole d'observation d'indices. Les particularités qui s'adresseraient à un seul instrument seront toutefois précisées.

Cette stratégie comporte trois volets. Le premier point considéré est celui de la construction des instruments d'observation. La proposition traite ensuite de l'encadrement de l'observateur par le biais du cadre de référence et de l'entraînement. Elle comporte enfin des indications sur les conditions de l'observation proposant des façons d'organiser la classe et deux modes d'échantillonnage de l'objet à observer.

1. La construction des instruments

Les techniques d'élaboration de la grille, de l'échelle et du protocole, ont été décrites en détail au chapitre précédent. Des critères de qualité pour leur construction ont été identifiés. Ces éléments ne sont pas présentés à nouveau dans ce chapitre-ci. Il faut rappeler cependant qu'ils sont essentiels et préalables à tout autre consigne visant la validité et la fidélité des résultats de l'observation. Il est important de recourir à ces techniques et de se référer à ces critères de qualité pour choisir ou construire des instruments d'observation appropriés aux besoins. Cette première démarche est déterminante pour la qualité des résultats de l'observation.

2. L'encadrement de l'observateur

La problématique définie fait référence à des situations qui sont vécues en classe régulière. L'observateur, c'est l'enseignant qui désire utiliser l'observation pour suivre le développement de diverses habiletés chez ses élèves.

2.1 Le cadre de référence

Pour être observable, le développement d'une habileté doit correspondre à une série de comportements prévisibles, attendus. Aussi, la première composante d'une stratégie de l'observation est de fournir à l'enseignant un cadre de référence au sujet de l'habileté qu'il veut observer.

Ce cadre donne la définition de l'habileté sous forme de comportements observables et apporte diverses informations sur le développement de cette habileté. Les sources de ce cadre de référence pourraient être, entre autres, une revue de littérature appropriée, un relevé d'informations provenant d'expériences faites dans le milieu scolaire, et un consensus établi par un certain nombre d'enseignants qui visent le développement de cette habileté par leur enseignement.

L'enseignant, guidé par le cadre de référence, qui pourrait être le même pour l'un ou l'autre des instruments d'observation retenus, doit acquérir une bonne connaissance de l'habileté qu'il désire observer, connaissance précédemment établie comme essentielle à la qualité des observations.

2.2 L'entraînement de l'observateur

Une deuxième composante de la stratégie proposée concerne l'entraînement des observateurs, c'est-à-dire les enseignants. L'entraînement pourrait avoir lieu soit à partir d'une situation réelle, pour les enseignants qui, à certaines périodes de la journée, n'ont pas charge d'élèves. Pour les enseignants du niveau primaire dont l'horaire de travail est continu, cet entraînement pourrait être fait à partir d'enregistrements sur film de situations vécues en classe.

Dans le premier cas, l'enseignant qui veut devenir observateur pourrait accompagner un enseignant qui serait un observateur d'expérience. Au moment où ce dernier, selon un plan établi, ferait figurer une séance d'observation à son horaire, il inviterait l'apprenti à venir observer les mêmes élèves, en même temps que lui. Par la suite, ils compareraient les résultats de leurs observations et tenteraient de résoudre les désaccords qui pourraient apparaître sur leurs fiches. L'apprenti pourrait également participer à l'organisation de la classe pour l'observation. Cet entraînement se poursuivrait jusqu'à ce qu'il y ait un bon degré d'accord entre les observateurs.

Ce degré d'accord ne s'établirait cependant que sur les données de deux observateurs. Si l'observateur d'expérience a développé quelque biais, l'apprenti pourra susciter une remise en question ou bien développer lui-même ce biais. C'est là un risque de ce mode d'entraînement qui par ailleurs pourrait être très fructueux par la possibilité qui serait donnée à l'apprenti d'être confronté aux multiples facettes de la pratique de l'observation en classe. Le risque de biais systématique, bien qu'en principe contrôlé par l'instrument utilisé, peut être considérablement réduit si, au début de la période d'entraînement, l'observateur d'expérience vérifie sa maîtrise en complétant un protocole d'observation à partir d'une situation enregistrée sur film et en comparant sa fiche d'observation avec la fiche-type pour ce vidéo.

Selon le degré d'accord obtenu, il pourrait accepter d'être tuteur ou décider de mettre à jour sa technique d'observation avant d'accepter ce rôle de tuteur.

On peut penser à une banque de films ou vidéo, reproduisant des situations réelles, montrant la mise en oeuvre de différentes habiletés par des groupes d'élèves de divers degrés de scolarité. L'entraînement d'un observateur pourrait consister, sous la surveillance d'un tuteur, à utiliser un instrument d'observation avec l'intention de parvenir à un bon degré d'accord avec la fiche-type de chaque vidéo que le tuteur aurait mis à son programme ou que l'enseignant aurait choisi selon ses priorités. Le tuteur aurait pour rôle de guider l'entraînement de l'observateur, et de l'aider à résoudre les désaccords qui peuvent survenir entre les observations faites par l'apprenti et celles figurant sur la fiche-type du même vidéo. Ce mode d'entraînement pourrait se révéler efficace et rapide, étant donné qu'il est possible de visionner plusieurs fois le même film ce qui permet à l'apprenti d'étudier ses erreurs et de reconnaître ses omissions et ses ajouts.

Un entraînement est nécessaire pour utiliser adéquatement l'un ou l'autre des trois instruments. La grille et l'échelle d'observation peuvent cependant requérir un entraînement moins long que celui qui précède l'emploi du protocole d'observation d'indices.

L'expérience de l'observateur est considéré comme un facteur de la qualité de ses observations. Aussi, lorsque, par l'entraînement, l'enseignant a atteint un niveau acceptable de justesse dans ses observations, on peut supposer qu'avec l'expérience, ce niveau est maintenu, sinon augmenté.

L'entraînement de l'observateur comporte aussi une information sur l'inférence et sur l'effet de halo de même qu'une mise en garde contre leurs conséquences. Enfin, on indique à l'enseignant les façons de réduire les effets observateur-observé.

3. Les conditions de l'observation

La situation à mettre en place pour une séance d'observation doit répondre à deux conditions. D'abord, qu'elle contienne, de façon aussi certaine que possible, les éléments recherchés, ensuite qu'elle soit représentative des situations habituelles vécues en classe.

Pour prendre le rôle d'observateur, l'enseignant pourrait toujours organiser sa classe pour que tous les élèves soient occupés à des activités individuelles, ou à des activités de groupe. Ces activités ne devraient pas requérir toute l'attention de l'enseignant. Il pourrait s'agir d'exercices à compléter, de réalisation de projets, de discussions par petits groupes, de travaux de laboratoire, de travaux d'équipe, etc. L'enseignant, s'il le juge à propos, pourrait informer ses élèves de ses intentions

d'observer le déroulement des activités de certains élèves ou de certains groupes, du but de cette observation et de la façon dont il entend procéder. Il ferait en sorte que la place qu'il occupe en classe de même que les instruments qu'il utilise ne nuisent pas aux activités des élèves, afin de minimiser chez ceux-ci les effets de la sensation d'être observés.

Les autres conditions de l'observation, tels l'horaire et la longueur de la séance d'obervation, seraient arrêtées en fonction de la nature du phénomène à observer de la possibilité de reproduire la même organisation de classe à quelques reprises, et de l'horaire des activités scolaires. Un horaire-type ne peut guère être proposé ici.

Par contre, deux types d'échantillonnage peuvent être suggérés. Ces échantillons sont nommés échantillons d'événements: l'échantillon un-zéro et l'échantillon instantané. Le premier convient au protocole d'observation d'indices alors que le deuxième peut servir de base aux trois types d'instruments.

3.1 L'échantillonnage un-zéro

Ce mode d'échantillonnage consiste à noter la présence ou l'absence d'un ou de quelques comportements qu'on veut observer, au cours de courtes périodes consécutives.

Ainsi, un enseignant habile avec le protocole d'observation d'indices qu'il désire utiliser et connaissant bien l'habileté qu'il va observer, choisirait comme l'a fait Anderson, quatre indices à observer chez cinq élèves à la fois. Il identifierait deux séries de six fiches de consignation. Le nom de chaque élève apparaît sur deux fiches, une pour chacune des deux séances d'observation prévues.

Au jour fixé, il organiserait les activités de sa classe de façon à être libéré de son activité d'animateur pour assumer le rôle d'observateur. A toutes les dix secondes, il consignerait ses observations au sujet de chacun des cinq élèves observés. La séance d'observation pourrait durer entre cinq et dix minutes pour un groupe de cinq élèves, selon la nature de l'habileté observée.

Au total, pour trente élèves, on pourrait compter entre trente et soixante minutes pour compléter une série de fiches de consignation. Le nombre de groupes d'élèves que l'enseignant observerait au cours d'une séance d'observation pourrait varier entre un seul et trois, selon l'âge des élèves et leur niveau d'autonomie, facteurs déterminant la durée d'un travail individuel, sans support de la part de l'enseignant.

Si le protocole d'observation d'indices pour l'habileté compte huit indices, il faut que l'enseignant planifie quatre séances d'observation, en préparant toujours deux séries de fiches de consignation pour chaque tournée d'observation. Il n'est pas proposé ici d'utiliser l'échantillonnage un-zéro avec un protocole comportant plus de huit indices au total.

Lorsqu'il a complété les deux ou quatre séries de fiches, l'enseignant procède à la compilation des données, en comptant pour chaque élève le nombre de fois que chaque indice a été noté un, c'est-à-dire présent, au cours de la durée totale des deux ou quatre séances d'observation. Cette compilation dégage un profil de l'habileté chez l'élève. L'enseignant peut évaluer ce profil en le comparant au profil attendu, décrivant la maîtrise de l'habileté observée. Ce profil modèle fait partie du cadre de référence que l'observateur connaît.

La somme de travail exigée par ce procédé pourrait se comparer avantageusement à celle demandée par un ou deux examens qu'il faut préparer et corriger. Le temps d'observation d'un maximum de deux ou quatre heures, selon le nombre d'indices du protocole, se comparerait également au temps requis pour l'administration d'un ou deux examens. De façon générale cependant, l'enseignant est plus actif pendant une séance d'observation que pendant la surveillance d'un examen.

Il faut souligner ici qu'en plus de présenter des caractéristiques comparables aux examens, l'observation possède des caractéristiques qui lui sont propres, ce qui la rend irremplaçable. Elle permet, entre autres, de recueillir des informations à propos des procédés mêmes et non pas seulement des résultats de l'application de ces procédés par les élèves.

Une telle utilisation du protocole d'observation d'indices avec échantillonnage un-zéro, tout en répondant adéquatement à l'ensemble des conditions d'ordre pratique posées par l'observation systématique des élèves en classe, pourrait fournir à l'enseignant des informations lui permettant de suivre le développement des habiletés chez ses élèves et d'orienter son action pour faciliter ce développement.

3.2 L'échantillonnage instantané

Ce mode d'échantillonnage amène l'observateur à porter son attention sur un élève seulement assez longtemps pour observer et enregistrer un comportement, puis à passer à l'observation d'un autre élève. On observe ainsi un grand nombre d'élèves en peu de temps, et on peut utiliser un instrument comportant jusqu'à vingt éléments ou indices à observer.

L'enseignant habile avec l'instrument d'observation qu'il désire utiliser, grille, échelle ou protocole d'observation, et ayant une bonne connaissance du cadre de référence de l'habileté qu'il veut observer, aurait inscrit le nom des élèves sur les fiches de consignation qui lui seront nécessaires. Il pourrait observer environ cinq élèves à la fois.

Dans le cas de la grille ou de l'échelle d'observation, les fiches de consignations reproduisent la liste des éléments à observer dans une première colonne et fournissent au moins cinq colonnes de cases en ligne avec les éléments. En haut de chacune de ces colonnes, un espace est réservé pour le nom de l'élève.

Les fiches de consignation du protocole d'indices sont élaborées selon le modèle illustré à la page 121 du présent document.

L'enseignant pourrait procéder à une ou deux séances d'environ dix minutes d'observation pour chaque groupe de cinq élèves. Le temps de classe ainsi utilisé serait d'une ou deux heures pour un groupe de trente élèves.

Ce procédé est plus économique que le précédent en termes de temps parce que le nombre d'éléments ou d'indices observés en une séance de dix minutes peut être près de trois fois plus élevé qu'avec l'échantillonnage un-zéro.

Le comportement qui ne correspond pas à l'un des quatre indices observés pendant le procédé un-zéro, peut être noté quand même par l'échantillon instantané, s'il correspond à l'un des nombreux éléments ou indices recherchés simultanément.

La compilation des résultats se fait selon le mode prévu par l'instrument utilisé. L'enseignant obtient, par ce procédé aussi, les informations qu'il désire pour suivre et accompagner le développement de certaines habiletés chez ses élèves.

La considération de la somme de travail exigée de l'enseignant et du temps de classe requis dans le cas de l'échantillonnage instantané, permet de conclure que ce mode d'utilisation de l'observation répond adéquatement à la situation problématique exposée au début.

L'enseignant pourrait donc utiliser les instruments qui emploient l'un ou l'autre de ces modes d'échantillonnage.

Tels sont donc les éléments de la stratégie d'observation systématique qui est proposée. Cette proposition comprend un ensemble d'indications au sujet de la fabrication des instruments, de l'encadrement de l'observateur, et des conditions de l'observation.

Bien qu'elle n'ait pas subi l'épreuve de la réalité par des expérimentations, cette stratégie semble applicable dans le cadre mentionné. Elle serait susceptible de faire figurer l'observation systématique des élèves en classe, au nombre des instruments de mesure utilisés par l'enseignant.

CONCLUSION

Cette recherche a tenté de rassembler, d'analyser et de présenter sous forme de synthèse, un ensemble de questions et de réponses concernant l'observation systématique.

Elle a aussi tenté d'adapter ces méthodes, conçues principalement en fonction des situations de recherche, à celles vécues en classe, dans le but de définir un moyen de mesure qui serait des plus utiles pour l'enseignant qui désire intervenir dans le développement des habiletés chez ses élèves.

La proposition qui est faite d'une stratégie de l'observation systématique pourrait être mise à l'essai. Les données de la réalité amèneraient sans doute des précisions qui font peut être défaut ici. Mais ce qui semble manquer le plus à ce stade-ci, ce sont les instruments d'observation eux-mêmes.

Les programmes d'enseignement donnent comme objectifs le développement de certaines habiletés. Il faudrait arriver à définir ces habiletés en termes d'indices et d'éléments observables. On pourrait amorcer cette tâche

de définition par une revue de littérature certes, mais aussi par l'écoute des enseignants qui auront observé ces habiletés, de façon systématique, avec seulement quelques indices en tête. On développerait ainsi, peu à peu, par consensus, des instruments valides et fidèles.

D'autre part, la formation des enseignants pourrait inclure des activités visant à les rendre capables d'utiliser l'observation systématique comme moyen de recueillir des informations à propos de leurs élèves.

L'observation systématique des élèves en classe pourrait ainsi se trouver au début et à la fin d'un long processus de recherche sur les habiletés dont l'école veut favoriser le développement.

REMERCIEMENTS

L'auteur tient à remercier monsieur Antonio Le Sieur, Ph. D., pour son aide précieuse à la réalisation de cette recherche. Elle veut aussi souligner le support que monsieur Michel Trahan, Ph. D., lui a donné.

BIBLIOGRAPHIE

ANDERSON, W.W., A Study of the Effects of Self-Directed Activity upon Quantity, Quality, and Variety of Responses in a Group Directed Reading-Thinking Activity, thèse de doctorat, non publiée, University of Virginia, 1971, cité par R.M. Brandt, "Studying Behavior in Natural Settings", New-York, Holt, 1972, 401 p.

BOYD, R.D., et M. VEREDEVAULT, The Observation and Recording of Behavior, in "Review of Educational Research", vol. XXXVI, no 5, décembre 1966, pp. 529-551

BRANDT, R.M., Studying Behavior in Natural Settings, New-York, Holt, 1972, 401 p.

CRONBACH, L.J., Essentials of Psychological Testing, 3e éd., New-York, Harper & Row, 1970, 650 p.

DeLandsheere, G., et V. DeLandsheere, Définir les objectifs de l'éducation, 2e éd., Paris, P.U.F., 1976, 293 p.

DELGADO, R.R., et J.M.R. DELGADO, An Objective Approach to Measurement of Behavior, in "Philosophy of Science", vol. 29, juillet 1962, pp. 253-268

DOWLEY, Edith M., Cues for Observing Children's Behavior, in "Childhood Education", vol. 45, no 9, mai 1969, pp. 517-521

ETAT DU MARYLAND, Montgomery County Public Schools, Early Identification of Learning Disabilities and Suggestions to Assist Classroom Teachers in Program Design, 1976, 72 p.

ETAT DE WASHINGTON, Bureau of Education for the Handicapped, Learning Accomplishment Profile (LAP), 1976, 135 p.

ETAT DE WASHINGTON, Bureau of Education for the Handicapped, Anne R. Sanford (Ed.), A Manual for Use of the Learning Accomplishment Profile, 1976, 245 p.

FLANDERS, N.A., The Problems of Observer Training and Reliability, in E. Amidon et J. Hough (Eds), "Interaction Analysis: Theory, Research and Application", Mass., Addison-Wesley, 1967

GARVEY, R., et T.R. KIMBALL, Development of a Perceptual-Motor Rating Scale for Classroom Teachers to Record Child Behavior, conférence donnée en mars 1978, 62e rencontre annuelle, American Educational Research Association, 35 p.

GRONLUND, Norman E., Measurement and Evaluation in Teaching, 3e éd., New-York, Macmillan, 1976, 590 p.

GHISELLI, E.E., et C.W. Brown, Personnel and Industrial Psychology, New-York, McGraw-Hill, 1955, pp. 87-125

GUILFORD, J.P., et B. FRUCHTER, Fundamental Statistics in Psychology and Education, New-York, McGraw-Hill, 1978, 545 p.

HERBERT, J., et C. ATTRIDGE, A Guide for Developers and Users of Observation Systems and Manuals, in "American Educational Research Journal", vol. 12, no 1, hiver 1975, pp. 1-20

HEYNS, R., et R. LIPPITT, Systematic Observational Techniques, in G. Lindzey (Ed.), "Handbook of Social Psychology", vol. 1, Cambridge, Addison-Wesley, 1954, pp. 370-404

HEYNS, R.W., et A.F. ZANDER, Observation of Group Behavior, in L. Festinger et D. Katz (Eds.), "Research Methods in the Behavioral Sciences", New-York, Holt, 1953, pp. 381-417

KERLINGER, F.N., Foundations of Behavioral Research, 2e éd., New-York, H.R.W., 1973, 741 p.

MASH, Eric J., et G. Makohoniuk, The Effects of Prior Information and Behavioral Predictability on Observer Accuracy, in "Child Development", vol. 46, no 2, 1975, pp. 19-513

MCGAW, B., J.L. WARDROP, et al... Classroom Observation Schemes: Where are the Errors?, in "American Educational Research Journal", vol. 9, no 1, pp. 13 à 27

MEDLEY, D.M., et H.E. MITZEL, Measuring Classroom Behavior by Systematic Observation, in N.L. Gage (Ed.), "Handbook of Research on Teaching", Chicago, Rand McNally, 1963, pp. 243-328

MORISSETTE, Dominique, Les examens de rendement scolaire, Québec, P.U.L., 1979, 390 p.

POPHAM, W. James, Educational Evaluation, Englewood Cliffs, Prentice-Hall, 1975, 328 p.

PROVINCE DE QUEBEC, Ministère de l'éducation, D.G.D.P., Programme d'étude -- primaire -- Mathématique, document de travail no 16-2300-00, Québec, Juin 1979

PROVINCE DE QUEBEC, Ministère de l'éducation, D.G.D.P., Programme d'étude -- secondaire -- Français langue maternelle, document no 16-3412-01, Québec, novembre 1979

REID, John B., Reliability Assessment of Observation Data: A Possible Methodological problem, in "Child Development", vol. 41, no 4, 1970, pp. 1143-1150

ROSENSHINE, B., Evaluation of Classroom Instruction, in "Review of Educational Research", vol. 40, 1970, pp. 279-300

ROSENSHINE, Barak, et N. FURST, The Use of Direct Observation to Study Teaching, in Robert M.W. Travers (Ed.), "Second Handbook of Research on Teaching", Chicago, Rand McNally, 1973, pp. 122-183

SELLTIZ, Claire, Lawrence S. WRIGHTSMAN, et al..., Les méthodes de recherche en sciences sociales, Montréal, H.R.W., 1977, 606 p., (Traduit par David Bélanger)

SMITH, R.L., C., McPhail, et al..., Reactivity to Systematic Observation with Film: a Field Experiment, in "Sociometry", vol. 38, no 4, 1975, pp. 536-550

STEVENS, S., Mathematics, Measurement and Psychophysics, in S. Stevens (Ed.), "Handbook of Experimental Psychology", New-York, Wiley, 1951

TENBRINK, T.D., Evaluation: A Practical Guide for Teachers, New-York, McGraw-Hill, 1974, 493 p.

TRAVERS, R.M.W., An Introduction of Educational Research, New-York, 1958, pp. 161-230

WEBB, E., D.T. CAMPBELL, et al..., Unobtrusive measures, Chicago, Rand McNally, 1966, 225 p.

WRIGHTSTONE, J.W., Observational Techniques, in C.W. Harris & M.R. Liba (Eds.), "Encyclopedia of Educational Research", 3e éd., New-York, Macmillan, 1960, pp. 927-933